

Aus der Klinik für Innere Medizin I, mit den Schwerpunkten Gastroenterologie,
Hepatologie, Pneumologie, internistische Intensivmedizin, Endokrinologie, Infektiologie,
Rheumatologie, Ernährungs-und Alterungsmedizin
Direktor: Prof. Dr. med. S. Schreiber
Sektion Rheumatologie, Leitung: Prof. Dr. med. B.F. Hoyer
im Universitätsklinikum Schleswig-Holstein, Campus Kiel
an der Christian-Albrechts-Universität zu Kiel

Organ Specific Self-Assessment (OSSA): ein Fragebogen zur organspezifischen Einschätzung des subjektiven Gesundheitszustandes

Inauguraldissertation
zur
Erlangung der Doktorwürde der Medizin
der Medizinischen Fakultät
der Christian-Albrechts-Universität zu Kiel

vorgelegt von

Ulrike Gsell

aus Kleve

Kiel 2019

1. Berichterstatter: Priv.-Doz. Dr. Rainald Zeuner, Klinik für Innere Medizin I mit den Schwerpunkten Gastroenterologie, Hepatologie, Ernährungs- und Altersmedizin

2. Berichterstatterin: Prof. Dr. Hanna Kaduszkiewicz, Institut für Allgemeinmedizin

Tag der mündlichen Prüfung: 09.10.2020

Zum Druck genehmigt, Kiel, den 28.08.2020

gez.: Prof. Dr. Alexander Arlt, Klinik für Innere Medizin I mit den

Schwerpunkten Gastroenterologie, Hepatologie, Ernährungs- und Altersmedizin

Inhaltsverzeichnis

1	Einleitung.....	1
1.1	Was ist Gesundheit ?.....	2
1.2	Fragebögen und VAS in der Medizin.....	3
1.3	Ziele der vorgelegten Arbeit.....	5
2	Material und Methode.....	6
2.1	Studiendesign.....	6
2.2	OSSA-Fragebogen.....	7
2.3	Untersuchung der Re-Test Variabilität.....	9
2.4	Patienten- und Arztdokumentation des OSSA-Fragebogens.....	10
2.5	Statistische Methoden.....	11
3	Ergebnisse.....	13
3.1	Untersuchung Re-Test-Variabilität.....	15
3.2	Einschätzung der allgemeinen Gesundheit.....	17
3.3	Einschätzung des Organsystem-spezifischen Gesundheitszustandes...20	
3.4	Übereinstimmung der Einschätzung von Patienten und Ärzten bezüglich der Hauptdiagnosedomäne.....	35
3.5	Daten der Support Vector Machine (SVM).....	38
3.6	Zusammenfassung der Ergebnisse.....	41
4	Diskussion.....	43
4.1	Diskussion der Methode.....	43
4.2	Diskussion der Ergebnisse.....	46
4.3	Zusammenfassung der Diskussion und Ausblick.....	49
5	Zusammenfassung.....	51
6	Literaturverzeichnis.....	53
7	Danksagung.....	58
8	Anhang.....	59

1 Einleitung

Eine hohe Zahl von ambulant und stationär versorgten Patienten weist mehr als zwei verschiedene chronische Erkrankungen auf und ist somit als multimorbide einzustufen. Mit dem Ansteigen des durchschnittlichen Lebensalters in den Industrieländern kommt es zu einer steigenden Prävalenz von chronischen Krankheiten.

Bislang gibt es kein standardisiertes Verfahren zur Messung der Multimorbidität. Aus diesem Grund fällt die Prävalenz je nach Studie sehr unterschiedlich aus.

In einer Literaturrecherche reicht die Prävalenz der Multimorbidität ab 65 Jahren von 55 bis 98 % (Marengoni, A., et al., 2011).

Multimorbidität stellt sowohl im Rahmen einer stationären Behandlung als auch bei einer ambulanten spezialärztlichen Behandlung Ärzte vor das Problem der korrekten Koordination und Wichtung der unterschiedlichen den Patienten belastenden Erkrankungen. Häufig sind dabei in Vorunterlagen nur einzelne Aspekte des Gesundheitszustandes des Patienten berücksichtigt.

In einer Studie wird darauf hingewiesen, dass erstmals in den Jahren 1984-86 im Rahmen der nationalen Gesundheitsstudien repräsentative Daten zum Thema subjektiver Gesundheitszustand vorgelegt wurden (Helmert, U., 2003).

In derselben Studie wird zudem folgendes Ergebnis beschrieben:

„In der vorliegenden Untersuchung hat sich gezeigt, dass die subjektive Einschätzung des eigenen Gesundheitszustandes anhand der simplen Frage ‚Wie würden Sie Ihren gegenwärtigen Gesundheitszustand einschätzen?‘, deren Beantwortung nur wenige Sekunden in Anspruch nimmt, als ein sehr guter Prädiktor für die Sterblichkeitsentwicklung angesehen werden muss.“ (Helmert, U., 2003).

Bereits in mehreren Studien wird hinreichend belegt, dass die subjektive Gesundheitseinschätzung ein guter Prädiktor für Mortalität ist (Carstensen, J., 2016).

So einfach die Einschätzung des Gesundheitszustandes wirkt, umso schwieriger gestaltet sich die Aussage, was mit der subjektiven Einschätzung des Gesundheitszustandes gemessen wird. Dies hängt mit der schwierigen Frage nach der Definition von Gesundheit zusammen.

1.1 Was ist Gesundheit ?

Das Wort Gesundheit stammt vom germanischen (ga)sunda ab, das so viel bedeutet wie stark und kräftig.

Das Spannungsfeld, das den Begriff Gesundheit umfasst, wird durch zwei Extrempole von Definitionen gut beschrieben:

Die wohl bekannteste Definition von Gesundheit stammt von der Weltgesundheitsorganisation:

„Gesundheit ist ein Zustand völligen psychischen, physischen und sozialen Wohlbefindens und nicht nur das Freisein von Krankheit und Gebrechen. Sich des bestmöglichen Gesundheitszustandes zu erfreuen ist ein Grundrecht jedes Menschen, ohne Unterschied der Rasse, der Religion, der politischen Überzeugung, der wirtschaftlichen oder sozialen Stellung.“ (Verfassung der Weltgesundheitsorganisation, 1948).

Würde man der WHO-Definition folgen, wäre inzwischen unter anderem aufgrund des demographischen Wandels, nur noch eine Minderheit der Weltbevölkerung als gesund einzuschätzen.

Demgegenüber beschreibt der Philosoph Friedrich Nietzsche einen komplett individualisierten Gesundheitsbegriff „Gesundheit ist dasjenige Maß an Krankheit, das es mir noch erlaubt, meinen wesentlichen Beschäftigungen nachzugehen.“

Und bemerkt weiter „Denn eine Gesundheit an sich gibt es nicht, und alle Versuche ein Ding derart zu definieren, sind kläglich missraten. Es kommt auf dein Ziel, deinen Horizont, deine Antriebe, deine Irrtümer und namentlich auf die Ideale und Phantasmen deiner Seele an, um zu bestimmen, was selbst für deinen Leib Gesundheit zu bedeuten habe.“ (Friedrich Nietzsche, 1887).

Trotz des Fehlens einer allgemeingültigen Gesundheitsdefinition wird versucht, die verschiedenen Gesundheitsaspekte von Erkrankungen in Fragebögen mit Aktivitätsskalen zu berücksichtigen. Dabei gewinnen Patienteneinschätzungen (patient reported outcome measures) eine zunehmende Bedeutung auch in der Evaluierung von Therapien (American Food and Drug Administration Guidance for Industry, 2009). Die Patienteneinschätzung in Form einer visuellen Analogskala stellt dabei ein häufiges Messinstrument dar.

1.2 Fragebögen und VAS in der Medizin

Erstmals wurde eine Visuelle Analogskala (VAS) 1921 von Hayes & Patterson als neue Methode zur Beurteilung von Arbeitern durch deren Vorgesetzte veröffentlicht.

Es handelt sich bei der VAS um ein kontinuierliches Messinstrument, bei dem die Befragten ihr Antwort auf einer Linie darstellen. Die konkrete Ausgestaltung obliegt dem Forscher, da keine feste Operationalisierung von VAS in der Literatur vorliegt (Funke, F., 2010).

In der Medizin werden Patientenskalen, die auf Selbstevaluation beruhen, als Goldstandard bei der Erfassung von Schmerz und Krankheitsaktivität eingesetzt (Miskowski C, et al., 2008).

In der Rheumatologie wird beispielsweise der DAS28 und der BASDAI eingesetzt, der zur Patienteneinschätzung der Krankheitsaktivität der Rheumatoider Arthritis

bzw. Spondylitis ankylosans dient, wobei in beiden eine visuelle Analogskala Anwendung findet.

Da multidimensionale Skalen mit deutlich Mehraufwand verbunden sind, sowie Verständlichkeit und Akzeptanz geringer sind, werden primär eindimensionale Skalen eingesetzt, unter diesen die VAS eine zentrale Rolle einnimmt (Litcher-Kelly, 2007; Eccleston, 1997; Huskisson, 1974; Katz, 1999).

Aufgrund des geringeren Aufwands können deutlich größere Patientenzahlen erfasst werden.

In einer Studie über den EuroQol-Fragebogen (EQ-5D) (zwischen 2002-2003), welcher ein Instrument zur subjektiven Beschreibung und Bewertung von Gesundheitszuständen ist und eine Analyse von soziodemografischen Einflussfaktoren liefert, wurde bereits das Thema 'subjektiver Gesundheitszustand' angesprochen. Auch hier wurde eine VA-Skala verwendet (König, H.H., et al., 2005).

Als genereller Fragebogen stellt der EQ-5D-5L somit ein gut validiertes Instrument für patient-reported outcomes dar.

Die Dokumentation der subjektiven Einschätzung des eigenen Gesundheitszustandes erfolgt regelmäßig mittels Visueller Analogskalen (VAS) und ist regelhafter Bestandteil der Dokumentation in klinischen Studien.

Zur Erfassung der subjektiven Gesundheit wurde 1998 bereits der SF-36 als Messinstrument eingesetzt (Morfeld, M., et al, 2008).

Allein eine Literaturanfrage am 01.06.2019 an die Datenbank PUBMED/NCBI ergab unter dem Stichwort „visual analogue scale“ 46765 Literaturverweise.

In vielen Studien wurde die visuelle Analogskala validiert und deren Nutzen bei der Erhebung von Daten belegt (Diener und Putzki, 2008) (vgl. Boge, 2015).

1.3 Ziele der vorgelegten Arbeit

Ziel der Arbeit ist es, einen kurzen Fragebogen zu entwickeln, der als patientenbasierte Auskunft das Gesundheitsempfinden bezüglich verschiedener Organsysteme erfasst. Hierfür wurde ein auf mehreren visuellen Analogskalen basierender Fragebogen konstruiert, in dem die Selbsteinschätzung des Gesundheitszustandes in Bezug auf verschiedene Organsysteme abgefragt wird (Organ-Specific-Self-Assessment). Der Fragebogen wird im Folgenden als „OSSA-Fragebogen“ bezeichnet.

Die vorliegende Arbeit soll als Pilotprojekt untersuchen, wie eine organsystem-spezifische Selbsteinschätzung des Patienten die Gesundheitssituation wiedergibt und ob sich die Patienteneinschätzung der jeweiligen Behandlung mit der Einschätzung der/ des behandelnden Ärztin/ Arztes korrelieren lässt.

Folgende Fragestellungen werden dabei untersucht:

- Wird „Gesundheit“ von dem Score abgebildet?
- Gelingt die Abbildung des Krankheitsempfindens des Patienten?
- Wird die Haupterkrankung von dem Fragebogen abgebildet?
- Werden relevante Nebenerkrankungen vom Score erfasst?
- Korrelieren die Einschätzungen von Ärzten und Patienten im OSSA-Fragebogen?

Das Projekt, das erste Daten zur Validierung des OSSA erheben soll, ist Teil einer weitergehenden Untersuchung, so dass zukünftig der Fragebogen den behandelnden Arzt in Entscheidungen der Ressourcenallokation (z.B. Erheben von fachspezifischen Anamnesen, bzw. weitergehenden Diagnostischen Maßnahmen) sinnvoll unterstützen kann.

2 Material und Methode

2.1 Studiendesign

Die Studie wurde als therapiebegleitende Untersuchung nach Überprüfung und zustimmender Bewertung durch die Ethikkommission des Universitätsklinikums Kiel gemäß der Deklaration von Helsinki durchgeführt. Nach vorheriger Aufklärung und schriftlicher Zustimmung erhielten die Probanden den standardisierten Fragebogen in anonymisierter Form. Die Daten wurden analog erhoben und in anonymisierter Form getrennt von der Krankenakte archiviert. Die Patientenrekrutierung erfolgte von November 2012 bis November 2015.

Die Auswahl der Probanden bezog sich auf viele Kliniken des UKSH Kiel, sowie der internistischen Notaufnahmen des UKSH, Campus Kiel und der internistischen Notaufnahme des Städtischen Krankenhauses Kiel. Zusätzlich wurden 50 Patienten aus der Allgemeinarztpraxis Dr. Lohmann in Preetz befragt. Die Patienten wurden aus vielen unterschiedlichen Instituten rekrutiert, um viele Patienten mit Erkrankungen verschiedener Organsysteme zu untersuchen.

Das Einschlusskriterium war eine ambulante oder stationäre Behandlung in einer der bezeichneten Institutionen, sowie das Fehlen von Ausschlusskriterien.

Ausgeschlossen waren Patienten

- im Alter unter 18 Jahre
- Patienten mit demenzieller Entwicklung
- Patienten, die aufgrund von Schmerzen o.ä. nicht in der Lage waren, einen Fragebogen ausfüllen zu können

Es wurden insgesamt 410 Patienten mit dem neu entwickelten Fragebogen befragt.

Folgende Institute/Fachbereiche nahmen an der Studie teil:

- Rheumatologie UKSH, Campus Kiel
- Lungenambulanz UKSH, Campus Kiel
- Klinik für Dermatologie UKSH, Campus Kiel
- Nephrologie UKSH, Campus Kiel
- Gastroenterologie UKSH, Campus Kiel
- zwei Notaufnahmen (Universitätsklinikum Schleswig-Holstein, Campus Kiel (= Notaufnahme 2) und internistische Aufnahmestation des städtischen Krankenhaus Kiel (= Notaufnahme 1))
- Kardiologie UKSH, Campus Kiel
- Praxis für Allgemeinmedizin
- Augenklinik UKSH, Campus Kiel
- Neurologie UKSH, Campus Kiel
- Adipositas-Sprechstunde der Ambulanz für Ernährungsmedizin UKSH, Campus Kiel

2.2 OSSA-Fragebogen

Der entwickelte Fragebogen erfasst zehn Fragen, welche sich auf zwei Seiten erstrecken. Die Patienten sollen den subjektiven Gesundheitszustand mit Hilfe einer visuellen Analogskalen einschätzen.

Dabei betrifft die erste Frage den allgemeinen Gesundheitszustand, die folgenden jeweils ein spezifisches Organsystem. Mit Hilfe des Fragebogens soll ein Gesamtbild und Überblick über Gesundheitsbeeinträchtigungen bezüglich verschiedener Organsysteme erfasst werden. Als Bezeichnung für den Fragebogen wurde das Akronym OSSA (=Organ-Specific-Self-Assessment) gewählt.

Die Fragen waren wie folgt formuliert:

- Wie empfinden Sie den allgemeinen Gesundheitszustand?
- Wie empfinden Sie den Gesundheitszustand von Herz- und Kreislauf?
- Wie empfinden Sie den Gesundheitszustand der Lungen?
- Wie empfinden Sie den Gesundheitszustand von Muskeln und Gelenken?
- Wie empfinden Sie den Gesundheitszustand von Magen, Darm und Verdauungssystem?
- Wie empfinden Sie den Gesundheitszustand des Stoffwechsels?
- Wie empfinden Sie den Gesundheitszustand von Nieren und Geschlechtsorganen?
- Wie empfinden Sie den Gesundheitszustand von Nerven und Gedächtnis?
- Wie empfinden Sie den Gesundheitszustand der Haut?
- Wie empfinden Sie den Gesundheitszustand von Augen und Ohren?

Seitlich neben jeder Frage ist eine visuelle Analogskala platziert, welche zehn Zentimeter lang ist und bei 0 %, 25 %, 50 %, 75 % und 100 % eine Markierung zur Orientierung besitzt. Die Patienten wurden gebeten, jede Frage mittels der VA-Skala

von "rechts = 100 % Gesundheit" bis "links = 0 % Gesundheit" zu bewerten und mit einem Kreuz oder Strich zu markieren.

Zudem wurden alle VA-Skalen gemäß den folgenden allgemeinen Regeln verankert: Die Patienten wurden angewiesen, ihren Gesundheitsstatus mit unter 75 % zu bezeichnen, wenn sie medizinische Maßnahmen (diagnostisch oder therapeutisch) als notwendig erachteten. Sie sollten den Gesundheitsstatus bei einem dringlichen Handlungsbedarf <50 % festsetzen und <25 % im Falle, dass der Patient sich als medizinischen Notfall einschätzt. Die OSSA-Fragebögen sind im Anhang gelistet. Zuvor hatten wir eine Testversion ohne diese Hilfestellungen in der Rheumatologie bei 9 Patienten getestet. Alle nachfolgenden Fragebögen sind mit diesen Zusätzen ausgestattet.

2.3 Untersuchung der Re-Test Variabilität

Zur Untersuchung der Re-Test-Variabilität wurde eine Patientengruppe gebeten den OSSA-Fragebogen mehrfach auszufüllen, wobei die klinische Situation zwischen den beiden Messungen keine Veränderung der klinischen Beschwerden erwarten ließ. Es wurden deshalb 29 Patienten aus der Rheumatologie, die sich einer mehrstündigen Infusionstherapie unterzogen, am selben Tag vor und nach der Behandlung untersucht. In diesem Zeitraum ist eine wirksame Veränderung des klinischen Zustandes nicht zu erwarten. Des Gleichen wurden 20 Patienten aus der Adipositas-Sprechstunde, die sich einer Formular-Diät-Behandlung unterzogen, im Abstand von zwei Wochen untersucht. Auch in dieser Situation ist innerhalb der beiden Messungen allenfalls eine minimale Änderung des klinischen Zustandes zu erwarten.

2.4 Patienten- und Arztdokumentation des OSSA-Fragebogens

Zur Bearbeitung der weiteren Fragestellungen und Beurteilungen wurden die behandelnden Ärzte gebeten, parallel zu dem Patientenfragebogen einen speziellen Arztfragebogen auszufüllen, der sich auf den Gesundheitszustand des Patienten beziehen soll. Dieser Fragebogen weist zusätzlich noch ergänzende Hinweise zur Normierung der Einschätzungen auf.

Diese sind wie folgt:

- bei Herz-und Kreislauf: Beispiele: NYHA 3: 30 %, stabil eingestellte arterielle Hypertonie ohne Komplikationen: 90 %, Belastungsdyspnoe beginnend: 75 %
- bei Lunge: Beispiele: COPD Gold 3: 20 %, Belastungsdyspnoe beginnend: 75 %, O2-Heimtherapie: 10 %
- Muskel und Gelenke: Beispiel: dokumentierbare Arthritiszeichen (Druckschmerz und Schwellung) von einem Gelenk als Erstmanifestation: 60 %
- Magen, Darm und Verdauungstrakt: Beispiele: klinisch signifikanter unbeabsichtigter Gewichtsverlust: 75 %, Diarrhöen mit beginnenden klinischen Exsikkosezeichen: 50 %
- Stoffwechsel: Beispiel: Diabetes mit chronischen Komplikationen (z.B. Fußsyndrom, Retinopathie): 50 %
- Niere und Geschlechtsorgane: Beispiele: Dialysepflichtigkeit: 10 %, symptomatischer Harnwegsinfekt mit Fieber: 70 %, chronische Anämie bei Menorrhagie: 70 %
- Nerven und Gedächtnis: Beispiele: Polyneuropathie mit beginnender Gehbehinderung: 70 %, Chronische Hemiparese mit Gehbehinderung 30 %, ohne Gehbehinderung 50 %
- Haut: Beispiele: Hautläsionen an sichtbaren Körperstellen die im Alltag auffallen: 70 %
- Auge und Ohren: Beispiele: Brillenträger ohne besonders starke Korrektur: 85 %, Schwerhörigkeit mit dauerhaftem Bedarf eines Hörgerätes: 50 %

Die Beurteilung sollte unter dem Gesichtspunkt eines weiteren Abklärungs- und Therapiebedarfs erfolgen. Auch hier gelten die gleichen Interventionsgrenzen bei 75 %, 50 % und 25 %.

Zusätzlich sollen, wenn vorhanden, die fünf wichtigsten Diagnosen (eine Hauptdiagnose und vier Nebendiagnosen) in der Reihenfolge ihrer Bedeutung für den Gesundheitszustand des Patienten wiedergegeben werden.

Da nicht von allen Patienten alle Fragen beantwortet wurden und auch der Rücklauf der Arzteinschätzungen nicht komplett waren, lagen 336 vollständige Datensätze mit Patientenfragebögen und Arztfragebögen für die weitere Auswertung vor.

2.5 Statistische Methoden

Die statistische Analyse wurde unter Verwendung nichtparametrischer Tests durchgeführt.

Der Wilcoxon-Rangsummentest und Kruskal-Wallis-Test wurden unter anderem als nicht-parametrische Testverfahren eingesetzt. Zusammenhänge zwischen Variablen wurden mit Hilfe der Spearman Korrelation analysiert. Zur Datenanalyse und -visualisierung sowie zur Modellbildung fanden die Programme Sigmastat 11 sowie R 3.6 Verwendung.

Das Signifikanzniveau wurde bei zweizeitiger Testung mit $p < 0,05$ festgelegt und bei Mehrfachtestung Bonferoni korrigiert.

Zur Vorhersage von Hauptdiagnosen basierend auf Patienten- und Arzteinschätzung wurde die Support Vector Machine (SVM) eingesetzt. Um die Modellleistung zu bewerten, wurde die Multiclass-AUC (Bereich unter der ROC-Kurve) gemäß (Hand and Till, 2001) anhand wiederholter Kreuzvalidierung (10-fach, 5 Wiederholungen) geschätzt, wobei die SVM-Hyperparameter mit Hilfe der Rastersuche optimiert wurde.

Die wiederholten Kreuzvalidierungen (10-fach, 5 Wiederholungen) wurde zur Bewertung der Klassifikationsgüte verwendet. Dies erlaubt eine Abschätzung der Verteilung des Gütemaßes.

„Die grundlegende Idee der Support Vector Klassifikation besteht darin, die Menge von Objekten durch eine Hyperebene in zwei Klassen zu unterteilen.“ (Fischer, J., 2007).

J. Fischer hebt folgende Vorteile der SVM hervor: Eine Klassifikation ist durch die Anwendung schnell möglich, da nicht der komplette Trainingsdatensatz benötigt wird. Zudem besitzt SVM eine hohe Generalisierungsfähigkeit, welche gut auf reale Probleme anzuwenden ist und das Arbeiten in hohen Dimensionen ist ebenfalls möglich.

Die Support Vector Machine ist demnach eine gute Methode um verschiedenste statistische Problematiken zu lösen.

ROC Kurven (receiver operating characteristic curve) sind ein wichtiges Werkzeug in der quantitativen Diagnostik und hilft die Genauigkeit von Vorhersagen zu bestimmen und wurde in den 1950er Jahren entwickelt. Die wichtigste Kennzahl ist die Fläche unter der Kurve AUC (Area under curve), welche einfach zu handhaben ist (Altenburg, H-P., 2009).

Eine ROC-Kurve trägt dazu Sensitivität (true positive rate, TPR) gegen 1-Spezifität (false positive rate, FPR) auf. Die Fläche unter der Kurve ist ein Maß für die Vorhersage-Güte. $AUC=1$ (Kurve verläuft entlang des linken und oberen Randes) entspricht einem perfekten Modell, $AUC=0.5$ (Kurve liegt im Mittel auf der Diagonalen) einem Zufallsmodell.

3 Ergebnisse

Teilgenommen haben 410 Patienten und deren behandelnde Ärzten, wobei insgesamt 527 Fragebögen verteilt wurden. In einigen Fragebögen wurden von den Patienten die Fragen nicht vollständig beantwortet (Rücklaufquote 68-100 % je nach Institut). Bei anderen Bögen lagen die Arzteinschätzungen nicht vollständig vor (Rücklaufquote 30-100 % je nach Institut).

336 Fragebögen (82 %) lagen in kompletter Form mit Patienteneinschätzung und Arzteinschätzung vor. Diese Daten wurden für die weiteren Analysen auch mittels Support Vector Machine verwendet.

Die Probanden wurden aus vielen verschiedenen Instituten ausgewählt, um Patienten mit Erkrankungen verschiedener Organsysteme zu untersuchen.

Die Fragebögen teilen sich in den Instituten wie folgt auf:

- Vorversuch mit 9 Patienten aus der Rheumatologie UKSH, Campus Kiel (OSSA ohne Orientierungshilfen)
- 20 aus der Rheumatologie UKSH, Campus Kiel (aktuelle Version des OSSA)
- 29 Patienten aus der Rheumatologie UKSH, Campus Kiel, mit einer zweiten longitudinalen Befragung
- 20 aus der Adipositassprechstunde der Ambulanz für Ernährungsmedizin UKSH, Campus Kiel, mit einer zweiten longitudinalen Befragung
- 52 Patienten aus der internistischen Notaufnahme 1 des städtischen Krankenhauses Kiel
- 50 Patienten aus der internistischen Notaufnahme 2 des UKSH, Campus Kiel
- 23 Patienten aus der Lungenambulanz des UKSH, Campus Kiel

- 20 Patienten aus der Klinik für Neurologie UKSH, Campus Kiel
- 21 Patienten aus der Klinik für Kardiologie UKSH, Campus Kiel
- 20 Patienten aus der Augenklinik UKSH, Campus Kiel
- 20 Patienten aus der Klinik für Dermatologie UKSH, Campus Kiel
- 36 Patienten aus der Klinik für Nephrologie UKSH, Campus Kiel
- 20 Patienten aus der Gastroenterologie UKSH, Campus Kiel
- 50 Patienten aus einer Praxis für Allgemeinmedizin

In Vorversuchen wurde an 9 Patienten aus der Rheumaambulanz zunächst eine Fragebogenversion mit visuellen Analogskalen ohne Verankerung/Orientierungshilfen untersucht. Hier zeigte sich eine so hohe Streuung der Ergebnisse, so dass diese rasch mit den Hilfestellungen verankert wurden.

Demnach wurden alle Fragebögen mit den folgenden Orientierungshilfen versehen: Die Patienten sollten einen Gesundheitszustand unter 75 % angeben, wenn nach Ansicht des Patienten medizinische Maßnahmen ergriffen werden sollten. Es sollte ein Gesundheitszustand <50 % angegeben werden, wenn nach Ansicht des Patienten dringlich medizinische Maßnahmen ergriffen werden sollten und <25 % sollte dokumentiert werden, wenn nach Ansicht des Patienten eine Notfallsituation vorliegt.

Die 9 Fragebögen ohne Verankerung wurden nicht in die weiteren Betrachtungen einbezogen.

Demnach konnten für die Berechnungen außerhalb der Support Vector Machine 401 Fragebögen verwendet werden.

3.1 Untersuchung Re-Test-Variabilität

Zur Untersuchung der Re-Test-Variabilität wurde eine Patientengruppe gebeten, den OSSA-Fragebogen zweifach auszufüllen, wobei die klinische Situation zwischen den beiden Messungen keine Veränderung der klinischen Beschwerden erwarten ließ. Es wurden deshalb 29 Patienten aus der Rheumatologie, die sich einer mehrstündigen Infusionstherapie unterzogen, am selben Tag vor- und nach der Behandlung untersucht. In diesem Zeitraum ist eine wirksame Veränderung des klinischen Zustandes nicht zu erwarten.

Des Gleichen wurden 20 Patienten aus der Adipositas-Sprechstunde, die sich einer Formular-Diät-Behandlung unterzogen, im Abstand von zwei Wochen untersucht. Auch in dieser Situation ist innerhalb der beiden Messungen allenfalls eine minimale Änderung des klinischen Zustandes zu erwarten.

In dem Rangsummen Vergleich (Wilcoxon-Mann-Whitney-Test) ergaben sich keine signifikanten Unterschiede zwischen den Messungen. Der Median aller Erstmessungen lag dabei bei 85 bei einem 25-75 % Intervall von 75-98 %. Der Median der Zweitmessungen lag bei 87 mit einem 25-75 % Intervall von 75-97 % ($p=0,414$; Abb1).

Die Pearson Korrelation mit $r=0,867$, $p<0,0001$ deutet auf eine gute Retest-Reliabilität des OSSA-Fragebogens hin.

Zur weiteren Beurteilung der Reliabilität wurde der Cronbach Alpha aus allen vollständig vorliegenden OSSA-Daten berechnet. Dieser liegt bei 0,87, was als Maß für eine gute interne Konsistenz angegeben wird (Nunnally, J.C., 1978).

Re-Test-Variabilität bei Patienten Rheumatologie und Adipositas n=49

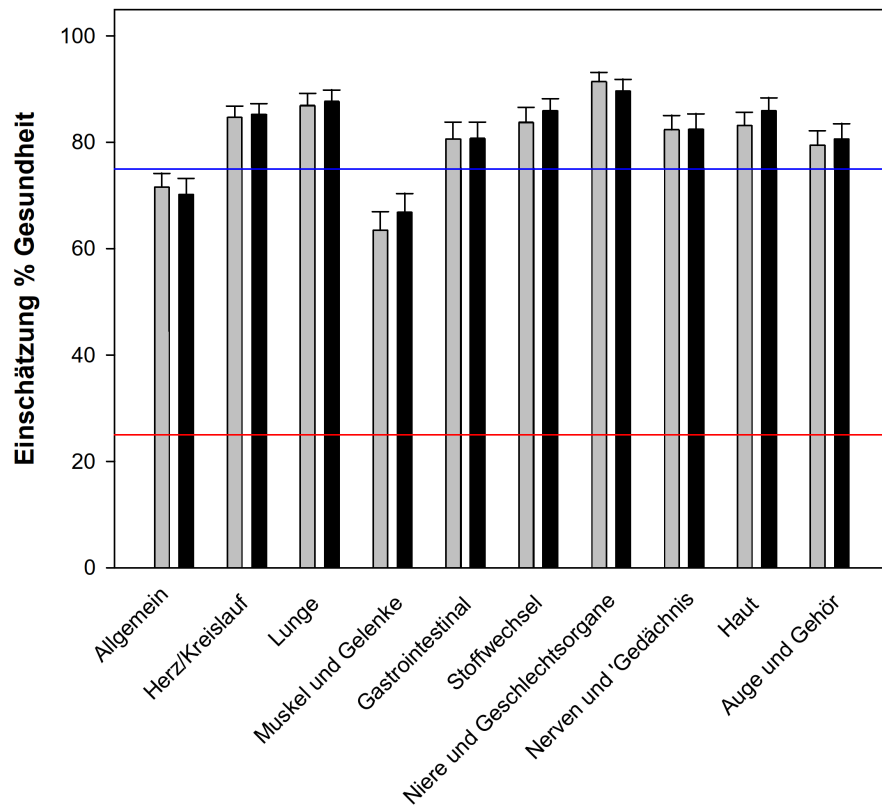


Abbildung 1: Re-Test-Variabilität bei 49 Patienten aus der Rheumatologie (n=29) und der Adipositas-Sprechstunde (n=20). Die erneute Testung (schwarze Balken) erfolgte in der Rheumatologie nach einer mehrstündigen rheumatologischen Infusionsbehandlung. In der Adipositas-Sprechstunde nach zwei Wochen. In beiden Situationen ist eine relevante Änderung des klinischen Befundes nicht zu erwarten. Dargestellt sind Mittelwerte und SEM.

3.2 Einschätzung der allgemeinen Gesundheit

Für die Einschätzung des allgemeinen Gesundheitszustandes gaben die Patienten einen Mittelwert von 66,9 % (SEM 1,18) an, der Median lag bei 75 %.

Die Einschätzung der Ärzte lag demgegenüber bei einem Mittelwert von 73,6 % \pm SEM 0,98, bei einem Median von 75 %.

Der Unterschied zwischen Patientenurteil und Arzturteil ist im Wilcoxon-Mann-Whitney Rangsummentest signifikant ($p < 0,001$; Abb.2). Dies bedeutet, dass Ärzte im Allgemeinen den Gesundheitszustand eher günstiger als die Patienten einschätzen. Die Mittelwerte und Mediane lagen dabei dicht an der für den Test vereinbarten Grenze von 75 % Gesundheit (Bedarf einer medizinischen Maßnahme).

Weiter wurde die Einschätzung des allgemeinen Gesundheitszustandes nach den Institutionen, in denen die Daten erhoben wurden, analysiert (Abb.3). Es zeigen sich zwei Ausnahmen von der oben genannten Beobachtung, dass die Arzteinschätzung in der Regel günstiger ist als die der Patienten. In der Adipositas-Sprechstunde und in der Nephrologie lagen die Arzteinschätzungen des allgemeinen Gesundheitszustandes unter denen der Patienten. Dies ist nachvollziehbar, da für beide Disziplinen bekannt ist, dass Patienten die klinischen Probleme unterschätzen.

Hinsichtlich der „Verankerungen“ des OSSA-Fragebogens (75 % für medizinischen Handlungsbedarf, 50 % für dringlichen Handlungsbedarf, 25 % für Notfall) zeigen sich erwartungsgemäß in der Notaufnahme 1 ($59,0 \pm \text{SEM } 3,23$) und Notaufnahme 2 ($57,1 \pm \text{SEM } 3,28$), sowie der Kardiologie ($58,8 \pm \text{SEM } 5,69$), die niedrigsten Gesundheitseinschätzungen, während die Patienten der Adipositas-Sprechstunde den besten allgemeinen Gesundheitszustand dokumentieren ($84,3 \pm \text{SEM } 2,96$).

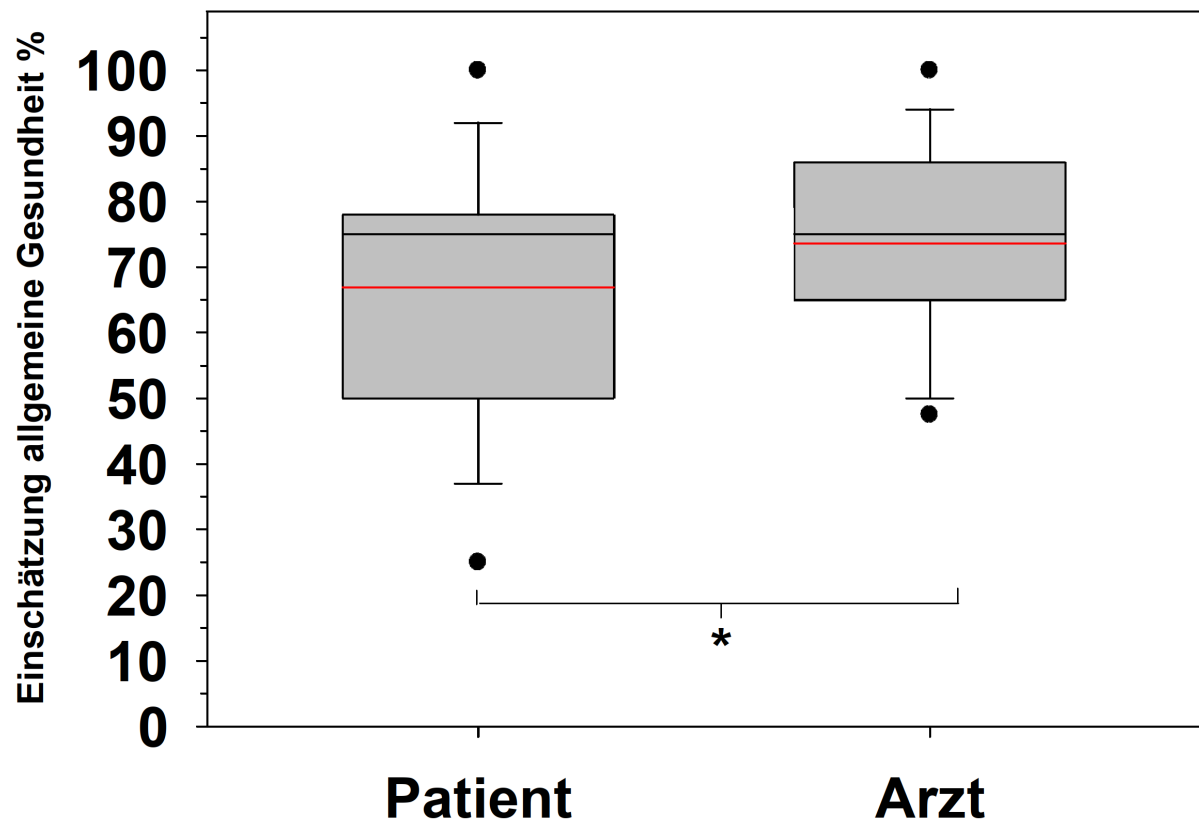


Abbildung 2: Darstellung der Einschätzung des allgemeinen Gesundheitszustandes seitens der Patienten und der Ärzte. Dargestellt sind 25 % und 75 % Häufigkeitsverteilung (Box) mit Median (schwarze Linie) und Mittelwert (rote Linie), sowie die 5 % und 95 % Intervalle und die Minimal- und Maximalwerte (Punkte). * = $p < 0,001$ im Mann-Whitney Rangsummentest.

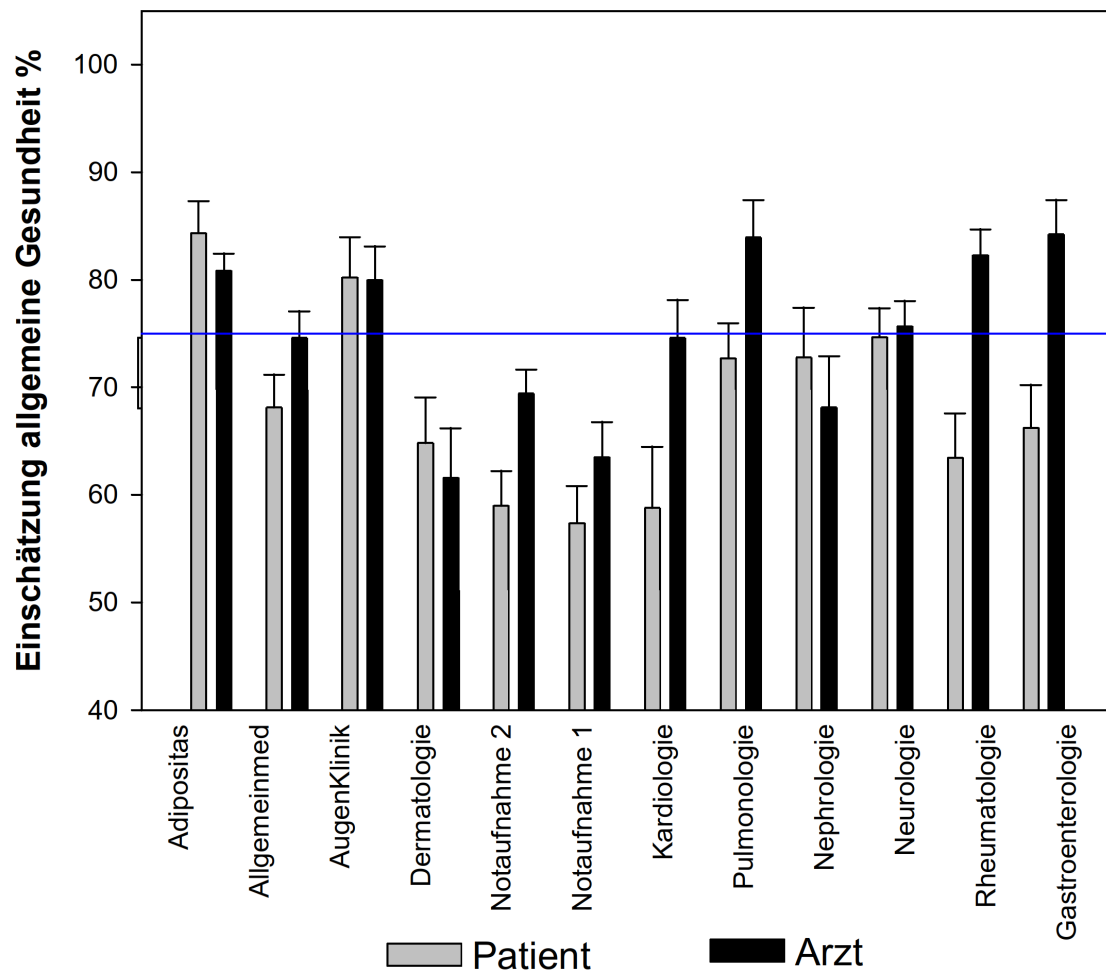


Abbildung 3: Einschätzung des allgemeinen Gesundheitszustandes im OSSA in verschiedenen Institutionen. Dargestellt sind Mittelwerte \pm SEM. Die grauen Balken bezeichnen die Patienteneinschätzung, die schwarzen Balken die Arzteinschätzung.

Tabelle 1: Einschätzung des allgemeinen Gesundheitszustandes von Patient und Arzt im Mittel je Klinik

Klinik	Einschätzung des allgemeinen Gesundheitszustandes der Patienten im Mittel	Einschätzung des allgemeinen Gesundheitszustandes der Ärzte im Mittel
Gastroenterologie	66,2 %	84,2 %
Allgemeinmedizin	66,9 %	74,6 %
Augenklinik	80,2 %	80,9 %
Dermatologie	64,8 %	61,6 %
Kardiologie	58,8 %	74,6 %
Nephrologie	72,3 %	66,5 %
Neurologie	74,7 %	75,6 %
Notaufnahme 1	59,1 %	69,3 %
Notaufnahme 2	57,1 %	62,8 %
Pulmologie	69,8 %	77,9 %
Adipositas	84,3 %	82,9 %
Rheumatologie	62,4 %	79,8 %

Mit einem mittleren Prozentsatz von $66 \pm 1,17$ SE bzw. $74 \pm 0,98$ SE war die OSSA-Bewertung der Ärzte demnach signifikant höher als die Bewertung der Patienten (p Wilcoxon $<0,001$).

3.3 Einschätzung des Organsystem-spezifischen Gesundheitszustandes

Betrachten wir die Ergebnisse zu den Organsystem-spezifischen Gesundheitszuständen. Hier wird die Zuordnungsfähigkeit der Patienteneinschätzung zu den jeweiligen Fachbereichen deutlich. Das Hauptproblem der Patienten lässt sich dem passenden Institut zuordnen. Die Ärzte bewerteten jeweils das gleiche Organsystem wie die Patienten als schlechtestes, sodass hier eine gute Übereinstimmung vorliegt.

Im Folgenden werden die jeweils schlechtesten Einschätzungen der Patienten und der Ärzte als jeweilige Hauptdiagnosedomäne bezeichnet. Als Nebendiagnosedomäne werden die zweit- schlechtesten Einschätzungen bezeichnet.

Tabelle 2: Auflistung der schlechtest bewerteten Organsysteme von Patient und Arzt im Mittel, geordnet nach Klinik

Fachbereich	Schlechtest bewertetes Organsystem des	
	Patienten im Mittel	Arztes im Mittel
Gastroenterologie	Magen-Darm 63 %	Magen-Darm 73,3 %
Allgemeinmedizin	Muskel&Gelenke 66,5 %	Muskel&Gelenke 80,2 %
Augenklinik	Auge&Ohr 57,4 %	Auge&Ohr 50,5 %
Dermatologie	Haut 51,8 %	Haut 53,3 %
Kardiologie	Herz&Kreislauf 60,5 %	Herz&Kreislauf 69,5 %
Nephrologie	Muskeln&Gelenke 71 %	Nieren 59,8 %
Neurologie	Nerven 67,7 %	Nerven 81,6 %
Notaufnahme 1	Herz&Kreislauf 67,0 %	Herz&Kreislauf 74,1 %
Notaufnahme 2	Herz&Kreislauf 65,5 %	Herz&Kreislauf 71.1 %
Pulmologie	Lunge 58,6 %	Lunge 67,0 %
Adipositas	Muskeln&Gelenke 73 %	Muskeln&Gelenke 84 %
Rheumatologie	Muskeln&Gelenke 57,6 %	Muskeln&Gelenke 74,4 %

Neben dem allgemeinen Gesundheitszustand lagen die Einschätzungen der Ärzte auch in der Hauptdiagnosedomäne höher als die der Patienten. Das heißt aus objektiver Sicht der Ärzte waren die Patienten gesünder als in der Selbsteinschätzung. Auch hier gilt Nephrologie als Ausnahme.

In folgenden Diagrammen werden die einzelnen Institute separat dargestellt. Hier werden Patienteneinschätzung und Arzteinschätzung direkt gegenübergestellt und man kann den Unterschied zwischen Selbsteinschätzung und Arzteinschätzung deutlich wahrnehmen.

POPGEN Organ Specific Self Assessment

Patienten Augenklinik n=20

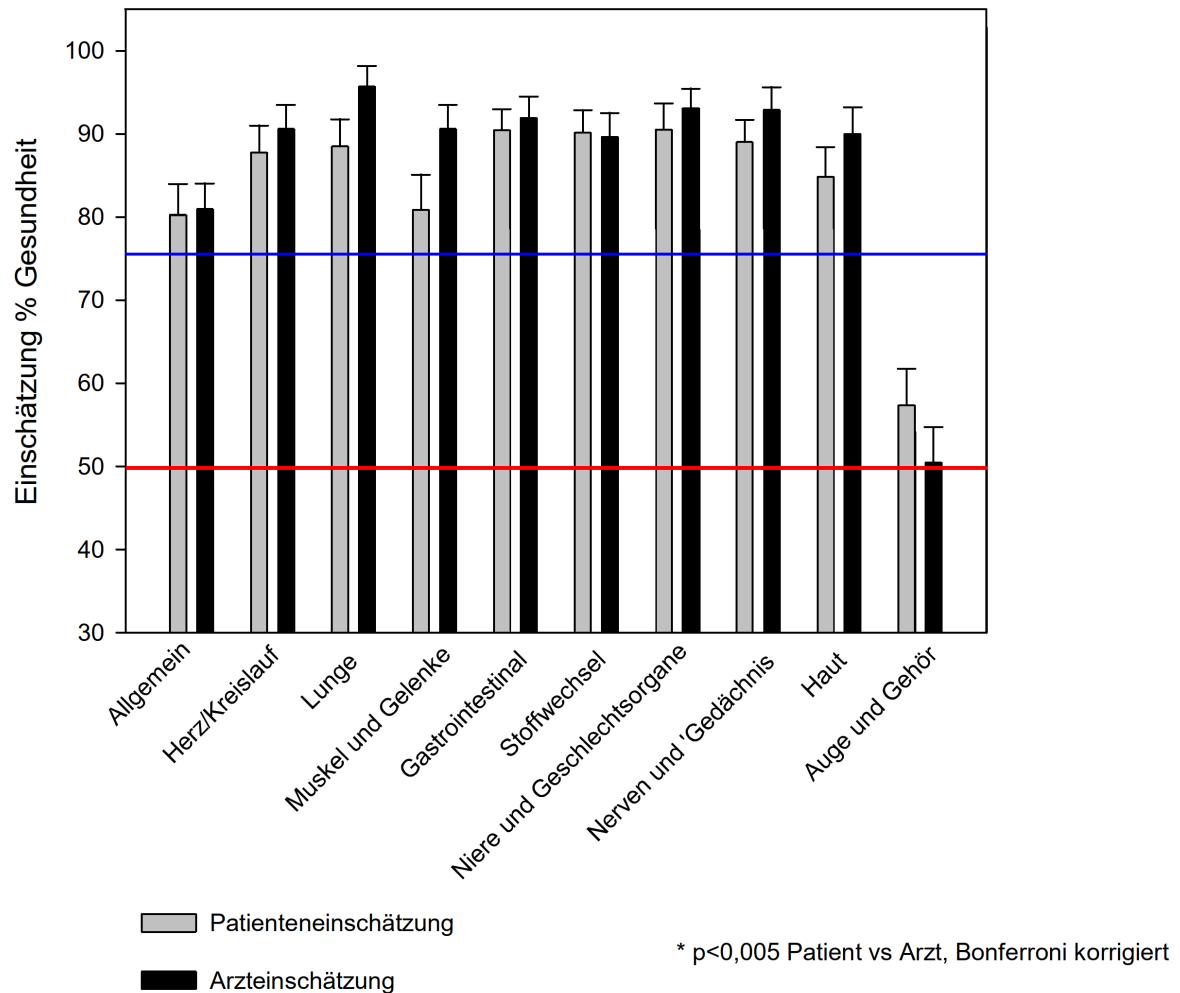


Abbildung 4: organspezifische Einschätzung in der Augenklinik, n=20; dargestellt als Mittelwerte \pm SEM; Patienteneinschätzung (grauer Balken) ist den Arzteinschätzungen (schwarzer Balken) direkt gegenübergestellt; der Cronbach alpha beträgt hier bei 0,87

Hier betrachten wir die organspezifische Einschätzung von Patienten und Ärzten in der Augenklinik. Hier wurden 20 Probanden und deren Ärzte mit dem OSSA befragt. Anhand Abbildung 4 kann man deutlich erkennen, dass das Organsystem „Augen und Ohren“ als schlechtestes bewertet wurden (Patient 57,4 %, Arzt 50,5 %). Der Cronbach alpha liegt hier bei 0,87.

POPGEN Organ Specific Self Assessment

Patienten Allgemeinmedizin n=50

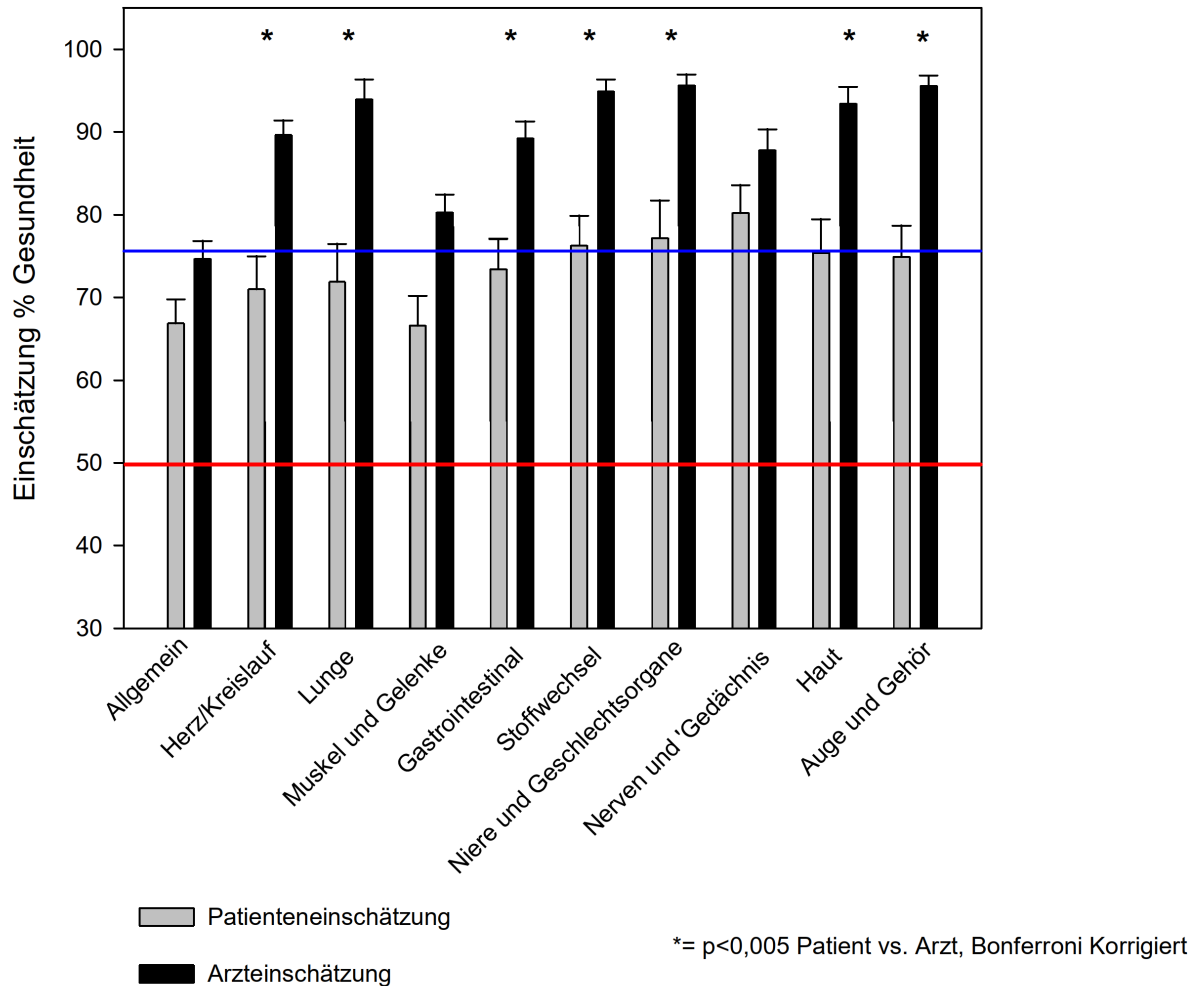


Abbildung 5: organspezifische Einschätzung in der Allgemeinmedizin, n=50; dargestellt als Mittelwerte±SEM; Patienteneinschätzung (grauer Balken) ist den Arzteinschätzungen (schwarzer Balken) direkt gegenübergestellt; der Cronbach alpha beträgt hier 0,92

In der allgemeinmedizinischen Praxis wurden 50 Probanden mit dem OSSA befragt. Hier wird Muskeln und Gelenke mit der niedrigsten Prozentzahl angegeben (Patient 66,5 %, Arzt 80,2 %), wobei hier eine breite Fächerung der Patienteneinschätzung bezüglich der Organsysteme vorliegt. Der Cronbach alpha liegt hier bei 0,92.

POPGEN Organ Specific Self Assessment

Patienten Dermatologie n=20

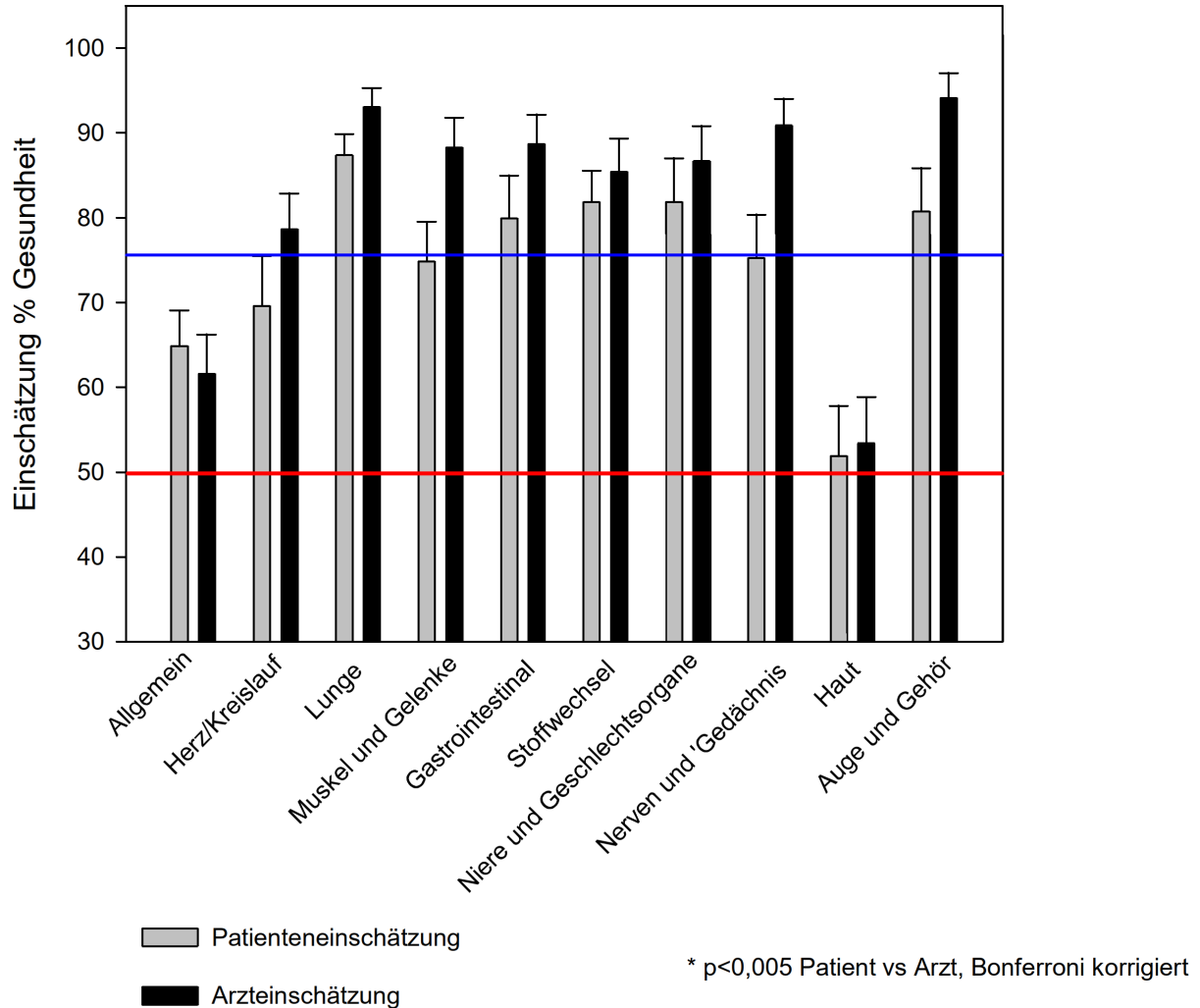


Abbildung 6: organspezifische Einschätzung in der Dermatologie, n=20; dargestellt als Mittelwerte \pm SEM; Patienteneinschätzung (grauer Balken) ist den Arzteinschätzungen (schwarzer Balken) direkt gegenübergestellt; der Cronbach alpha liegt hier bei 0,78

20 Probanden sollten in der Dermatologie ihren Gesundheitszustand mittels OSSA einschätzen. Das Organsystem „Haut“ zeigt auffällig niedrige Werte auf beiden Seiten (Patient 51,8 %, Arzt 53,8 %). Der Cronbach alpha liegt hier bei 0,78.

POPGEN Organ Specific Self Assessment

Patienten Gastroenterologie n=20

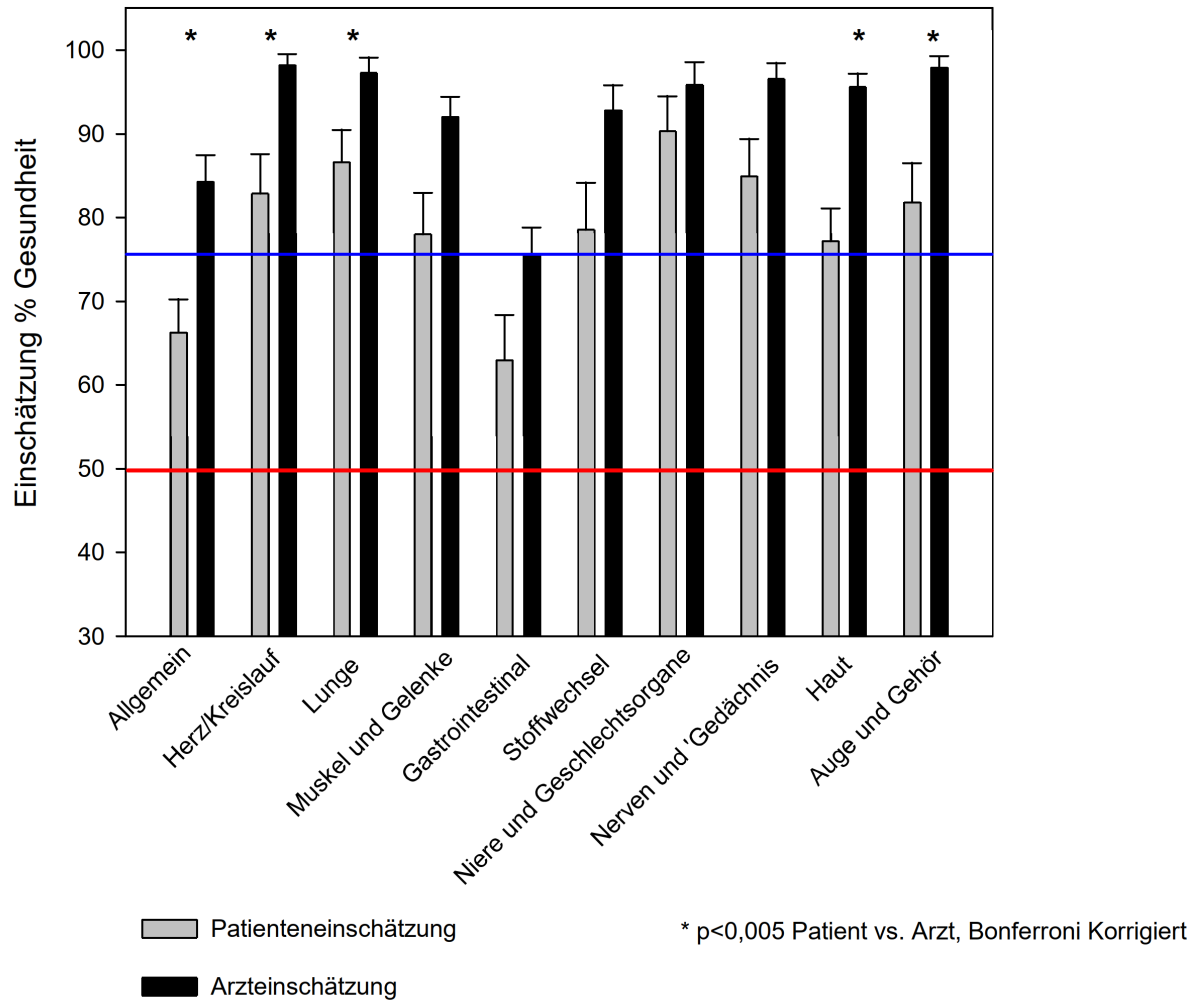


Abbildung 7: organspezifische Einschätzung in der Gastroenterologie, n=20; dargestellt als Mittelwerte±SEM; Patienteneinschätzung (grauer Balken) ist den Arzteinschätzungen (schwarzer Balken) direkt gegenübergestellt, Cronbach alpha liegt hier bei 0,92

Betrachten wir die Gastroenterologie mit 20 ausgefüllten OSSA-Fragebögen. Auch hier ist passend Magen-und Darm als schlechtestes eingeschätzt worden. Man kann einen deutlichen Unterschied zwischen Patienteneinschätzung (63 %) und Arzteinschätzung (73,3 %) erkennen. Hier liegt der Cronbach alpha bei 0,92.

POPGEN Organ Specific Self Assessment

Patienten Kardiologie n=21

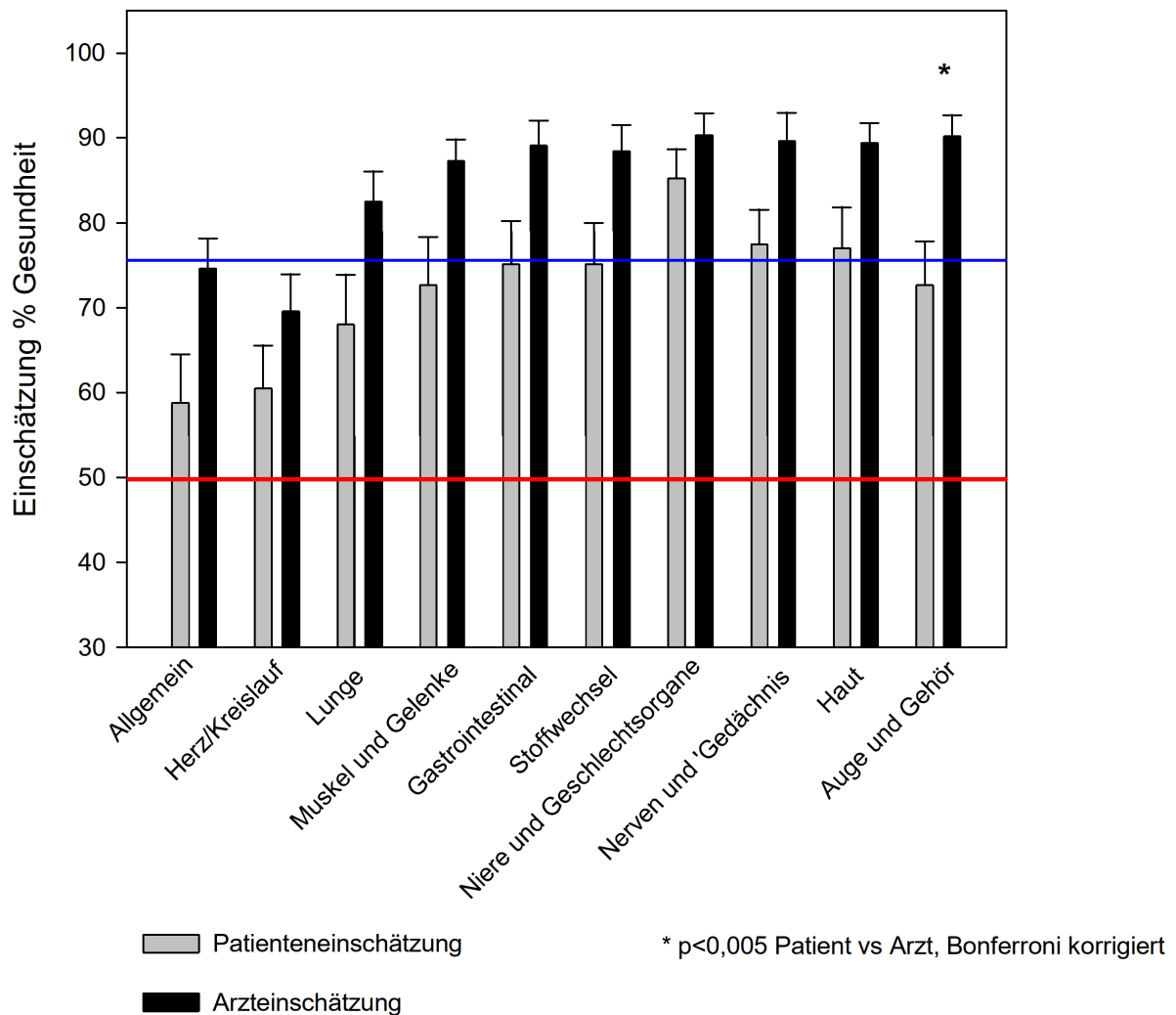


Abbildung 8: organspezifische Einschätzung in der Kardiologie, n=21; dargestellt als Mittelwerte \pm SEM; Patienteneinschätzung (grauer Balken) ist den Arzteinschätzungen (schwarzer Balken) direkt gegenübergestellt; der Cronbach alpha liegt hier bei 0,93

21 Patienten und deren Ärzte füllten den OSSA-Fragebogen in der Kardiologie aus. Herz-und Kreislauf wurden sowohl beim Patienten (60,5 %), als auch beim Arzt (69,5 %) mit der niedrigsten Einschätzung bewertet. 0,93 ist hier der Cronbach alpha.

POPGEN Organ Specific Self Assessment

Patienten Nephrologie n=36

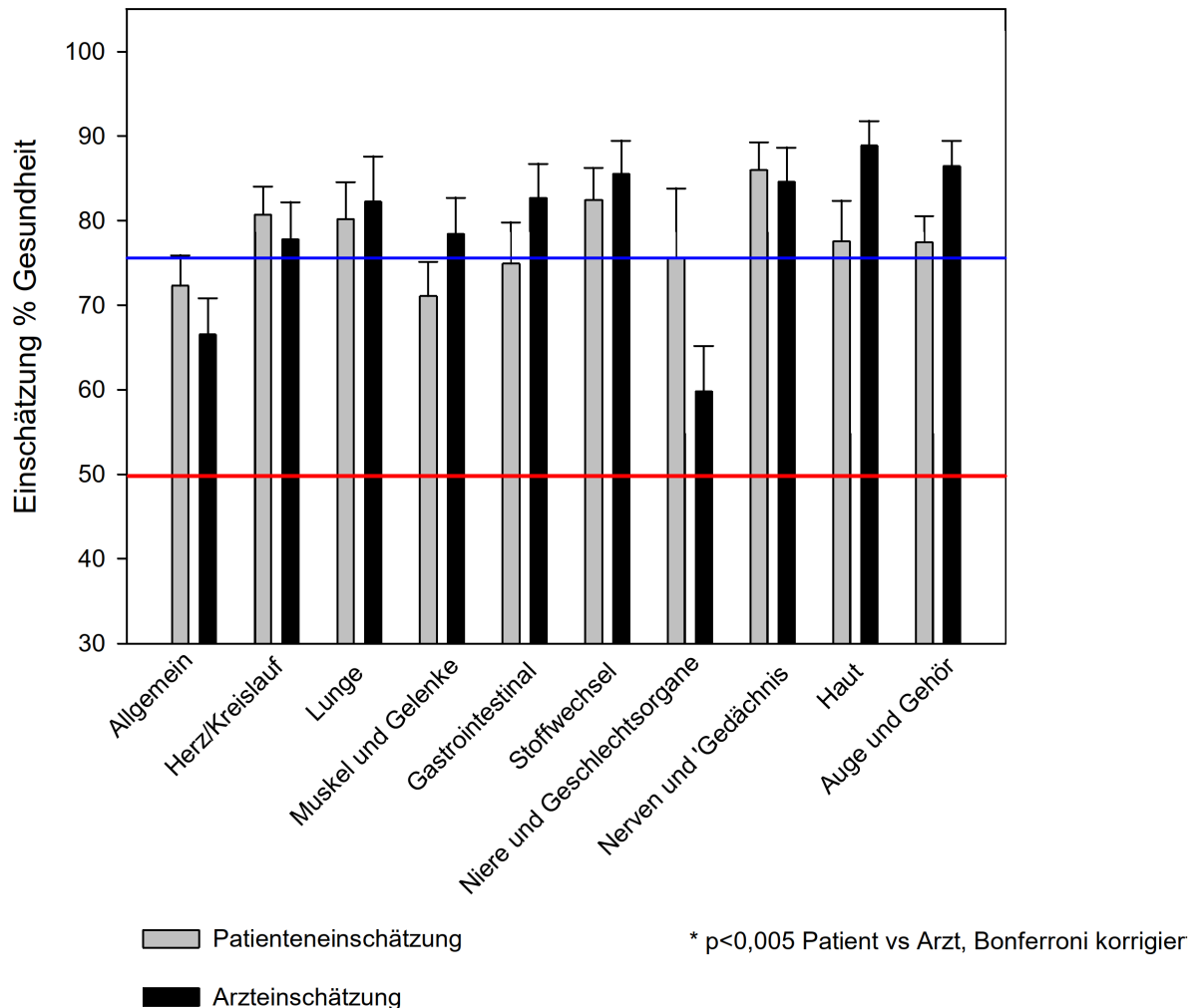


Abbildung 9: organspezifische Einschätzung in der Nephrologie, n=36; dargestellt als Mittelwerte \pm SEM; Patienteneinschätzung (grauer Balken) ist den Arzteinschätzungen (schwarzer Balken) direkt gegenübergestellt; der Cronbach alpha liegt hier bei 0,83

In der Nephrologie erfolgte die Befragung bei 36 Patienten. Das Item „Nieren und Geschlechtsorgane“ wurde von den Ärzten als schlechtestes eingeschätzt (59,8 %), von den Patienten wurde Niere nur als zweit schlechtestes angegeben (75 %). Hier wird die Ausnahme deutlich, dass die Patienten sich gesünder einschätzten als die Ärzte. Der Cronbach alpha liegt hier bei 0,83.

POPGEN Organ Specific Self Assessment

Patienten Neurologie n=20

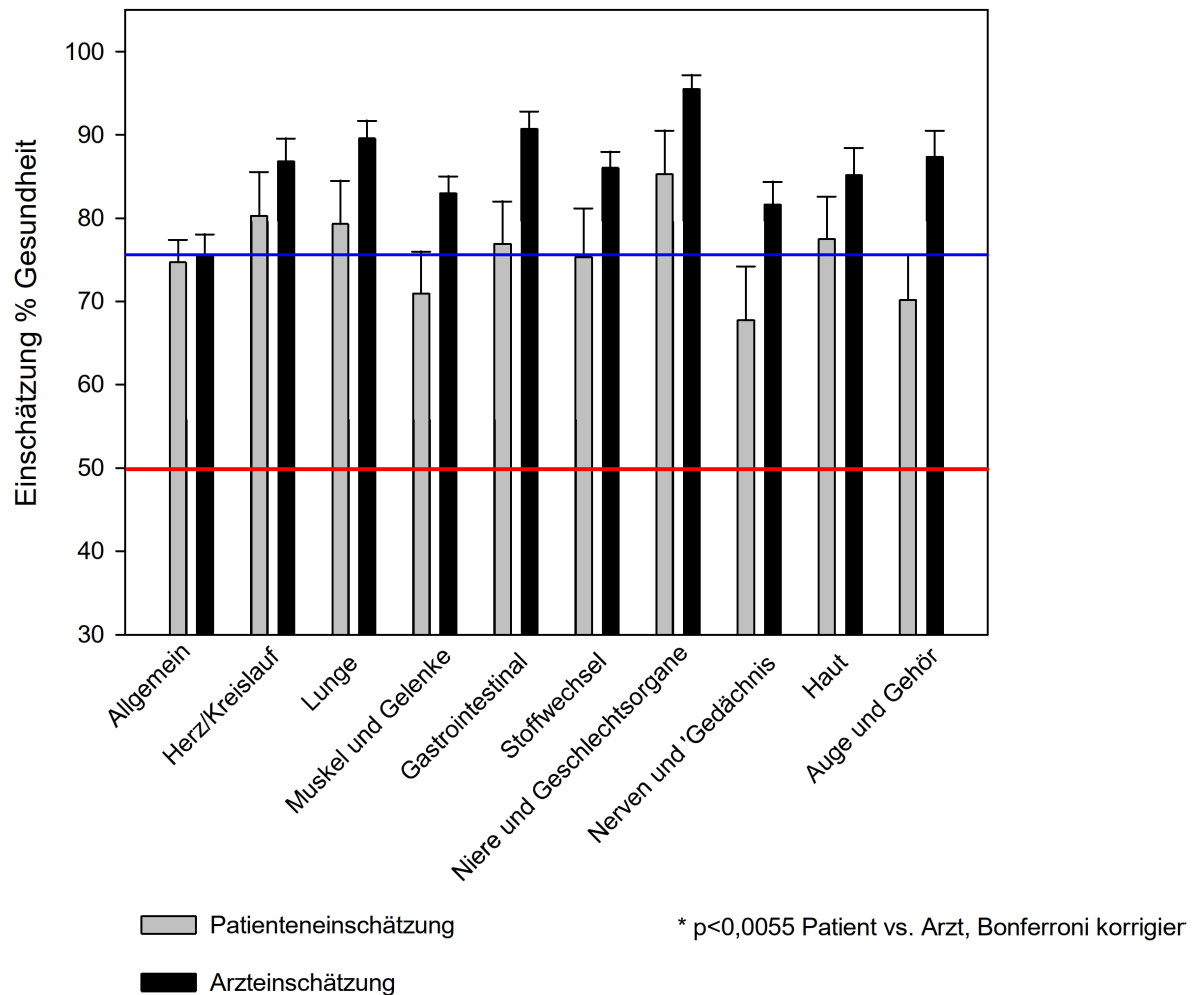


Abbildung 10: organspezifische Einschätzung in der Neurologie, n=20; dargestellt als Mittelwerte \pm SEM; Patienteneinschätzung (grauer Balken) ist den Arzteinschätzungen (schwarzer Balken) direkt gegenübergestellt; der Cronbach alpha liegt hier bei 0,91

Die 20 befragten Patienten und Ärzte in der Neurologie bewerteten das Organsystem „Nerven und Gedächtnis“ am niedrigsten (Patient 67,7 %, Arzt 81,6 %). Hier liegt der Cronbach alpha bei 0,91.

POPGEN Organ Specific Self Assessment

Patienten Notaufnahme 1 n=52

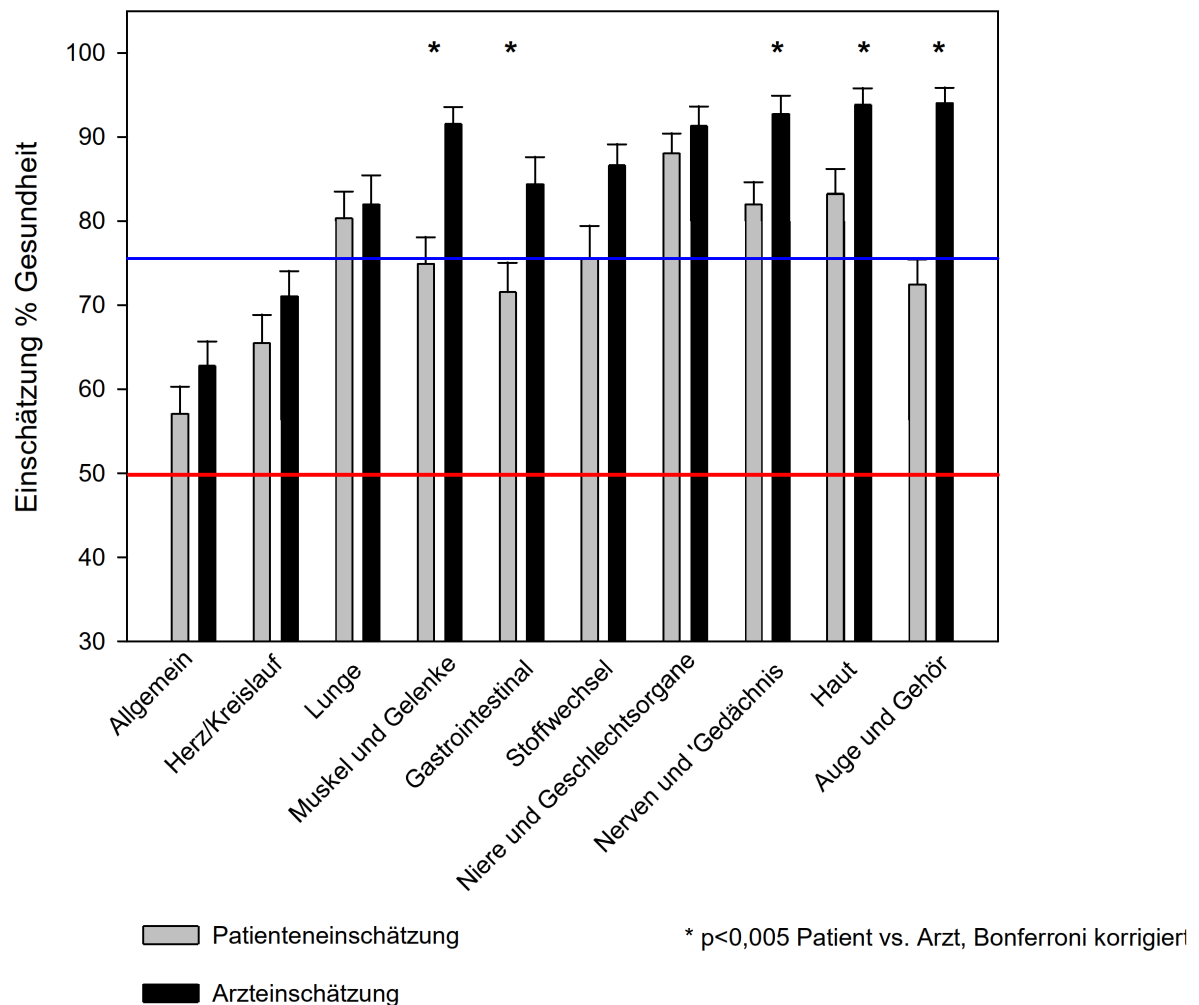


Abbildung 11: organspezifische Einschätzung in der Notaufnahme 1, n=52; dargestellt als Mittelwerte \pm SEM; Patienteneinschätzung (grauer Balken) ist den Arzteinschätzungen (schwarzer Balken) direkt gegenübergestellt; der Cronbach alpha liegt hier bei 0,88

In der Notaufnahme 1 erfolgte die Befragung bei 52 Probanden. Auch hier zeigt sich eine breite Fächerung, wobei sowohl beim Patienten (67,0 %), als auch beim Arzt (74,1 %) Herz und Kreislauf als schlechtestes bewertet wurde. 0,88 ist hier der Cronbach alpha.

POPGEN Organ Specific Self Assessment

Patienten Notaufnahme 2 n=50

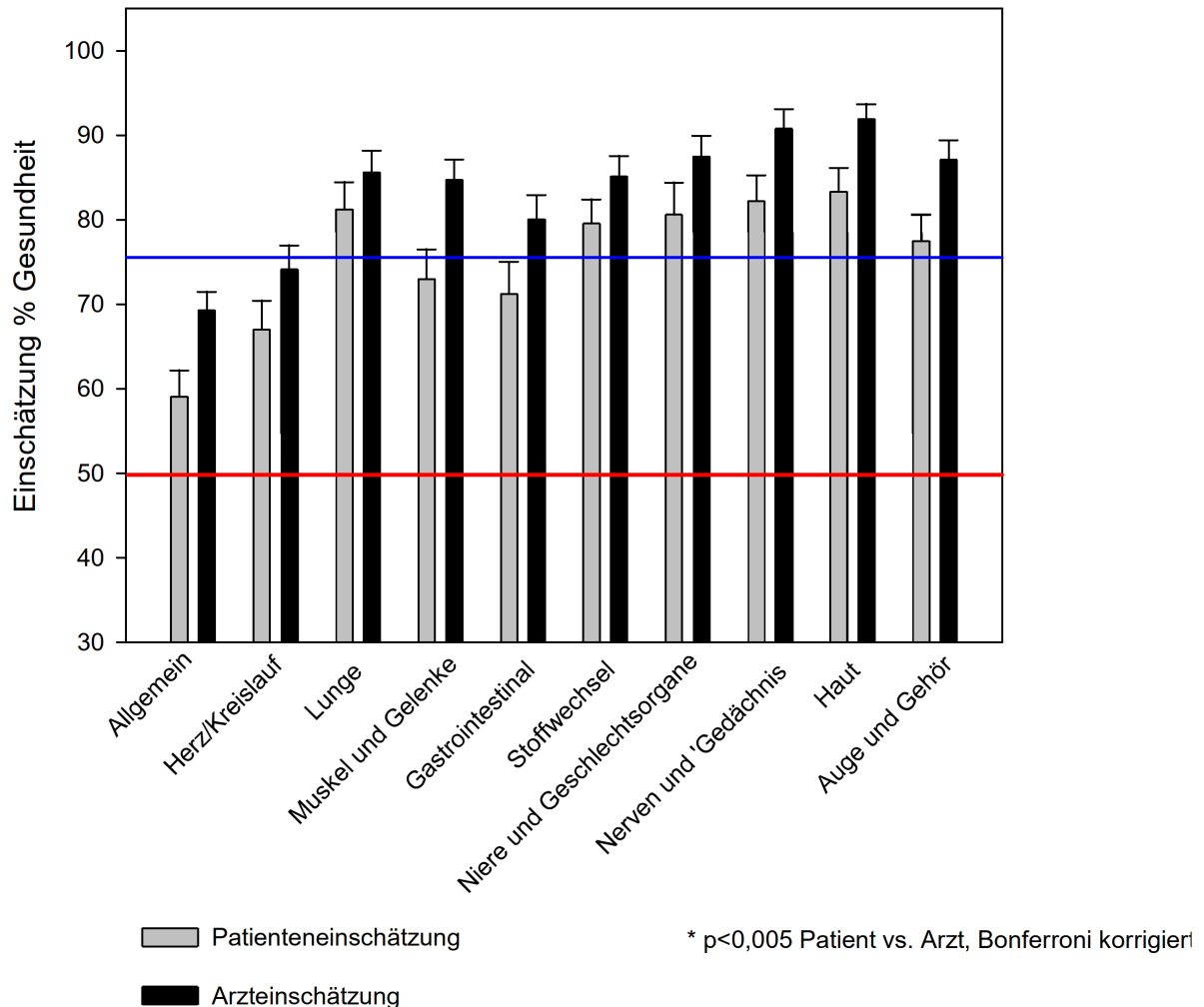


Abbildung 12: organspezifische Einschätzung in der Notaufnahme 2, n=50; dargestellt als Mittelwerte \pm SEM; Patienteneinschätzung (grauer Balken) ist den Arzteinschätzungen (schwarzer Balken) direkt gegenübergestellt; der Cronbach alpha liegt hier bei 0,8

Auch in der Notaufnahme 2 bei 50 Probanden zeigt sich ebenfalls Herz und Kreislauf mit niedrigsten Prozentwerten (Patient 65,5 %, Arzt 71,1 %). Das Verteilungsmuster der beiden Notaufnahmen zeigt eine erstaunliche Übereinstimmung. Der Cronbach alpha liegt hier bei 0,8.

POPGEN Organ Specific Self Assessment

Patienten Adipositassprechstunde n=20

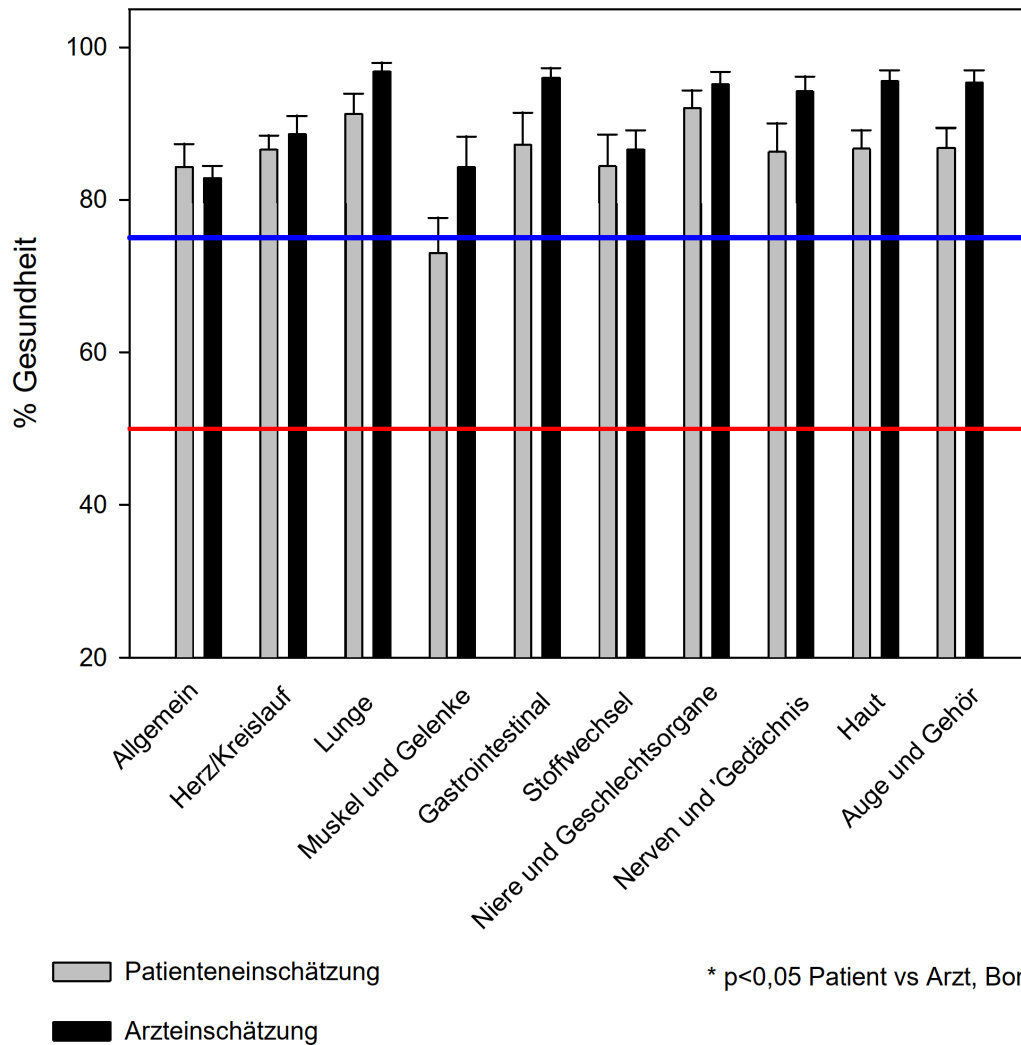


Abbildung 13: organspezifische Einschätzung in der Adipositas-Sprechstunde, n=20; dargestellt als Mittelwerte \pm SEM; Patienteneinschätzung (grauer Balken) ist den Arzteinschätzungen (schwarzer Balken) direkt gegenübergestellt; der Cronbach alpha liegt hier bei 0,84

20 Patienten wurden in der Adipositassprechstunde mit dem OSSA befragt. Hier wurde von den Patienten und den Ärzten das Organsystem „Muskeln und Gelenke“ (Patient 73 %, Arzt 84 %) statt „Stoffwechsel“ (Patient 84 %, Arzt 86 %) als schlechtestes bewertet. Hier mit einem Cronbach alpha von 0,84.

POPGEN Organ Specific Self Assessment

Patienten Pulmonologie n=23

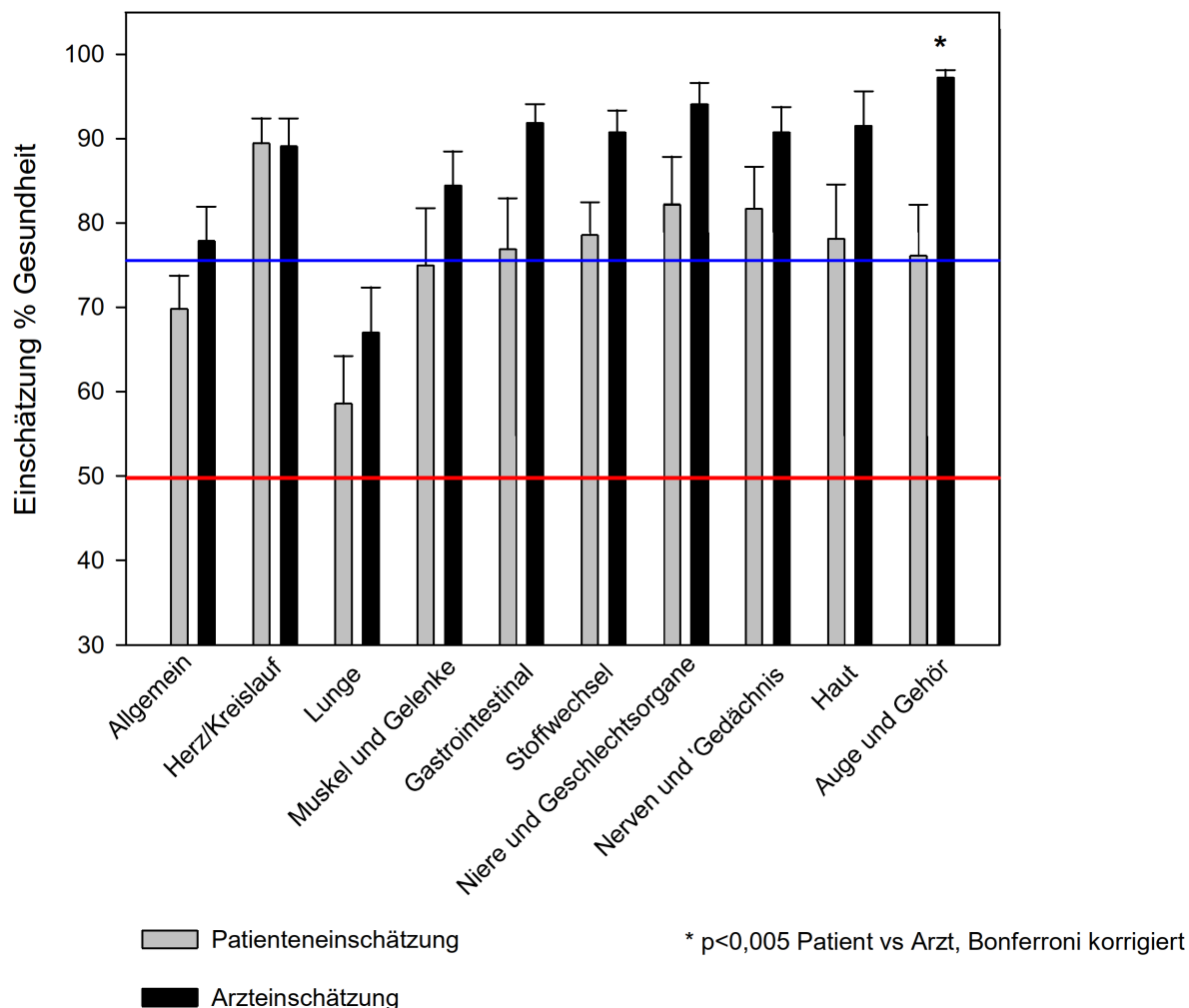


Abbildung 14: organspezifische Einschätzung in der Pulmologie, n=23; dargestellt als Mittelwerte \pm SEM; Patienteneinschätzung (grauer Balken) ist den Arzteinschätzungen (schwarzer Balken) direkt gegenübergestellt; der Cronbach alpha liegt hier bei 0,92

In der Lungenambulanz erfolgte die Befragung von 23 Patienten und deren behandelnden Ärzte. Auch hier stellt sich passend das Item „Lunge“ als schlechtestes bewertetes dar und ebenfalls schätzten sich die Patienten schlechter ein als die Ärzte (Patient 58,6 %, Arzt 67 %). 0,92 berechnet sich hier der Cronbach alpha.

POPGEN Organ Specific Self Assessment

Patienten Rheumatologie n=20

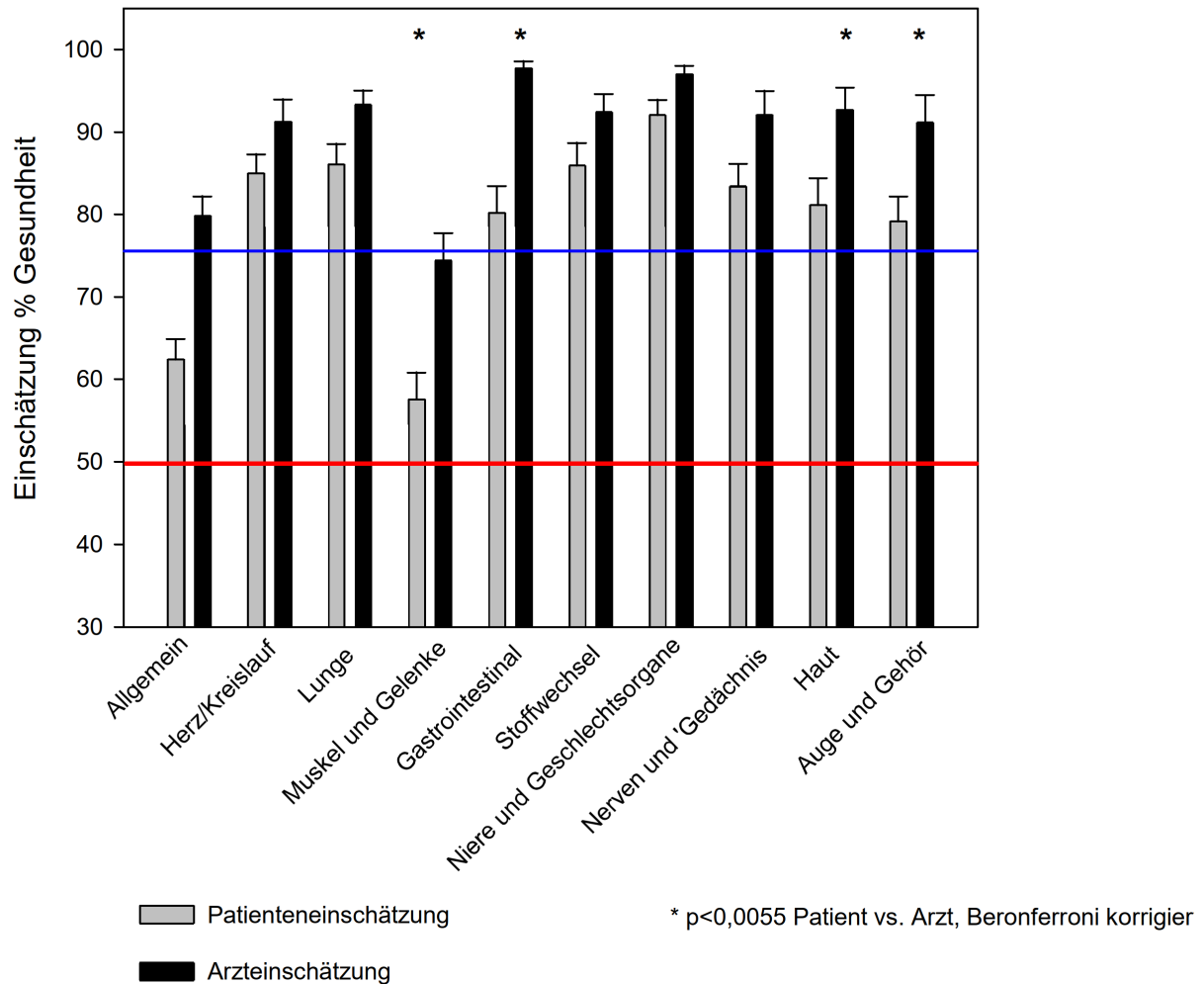


Abbildung 15: organspezifische Einschätzung in der Rheumatologie, n=20; dargestellt als Mittelwerte±SEM; Patienteneinschätzung (grauer Balken) ist den Arzteinschätzungen (schwarzer Balken) direkt gegenübergestellt; hier liegt der Cronbach alpha bei 0,8

Die 20 Probanden aus der Rheumatologie bewerteten „Muskeln und Gelenke“ mit der niedrigsten Prozentzahl (Patient 57,6 %, Arzt 74,4 %). Der Cronbach alpha liegt hier bei 0,8.

In der direkten Gegenüberstellung von Patienten- und Arzteinschätzungen wird besonders deutlich, dass die Patienten sich nicht nur wie bereits erwähnt im Item „allgemeine Gesundheit“ schlechter einschätzen, sondern in nahezu allen Organsystemen.

Die Hauptdiagnosedomäne führt zu den niedrigsten Einschätzungswerten in den entsprechend der Organsystemdomänen. Der Fragebogen erlaubt also eine korrekte Zuordnung des klinischen Problems des Patienten zu dem betroffenen Organsystem.

Abbildung 16 zeigt die Häufigkeitsverteilung der niedrigsten Gesundheitseinschätzungen in den verschiedenen Fachdisziplinen.

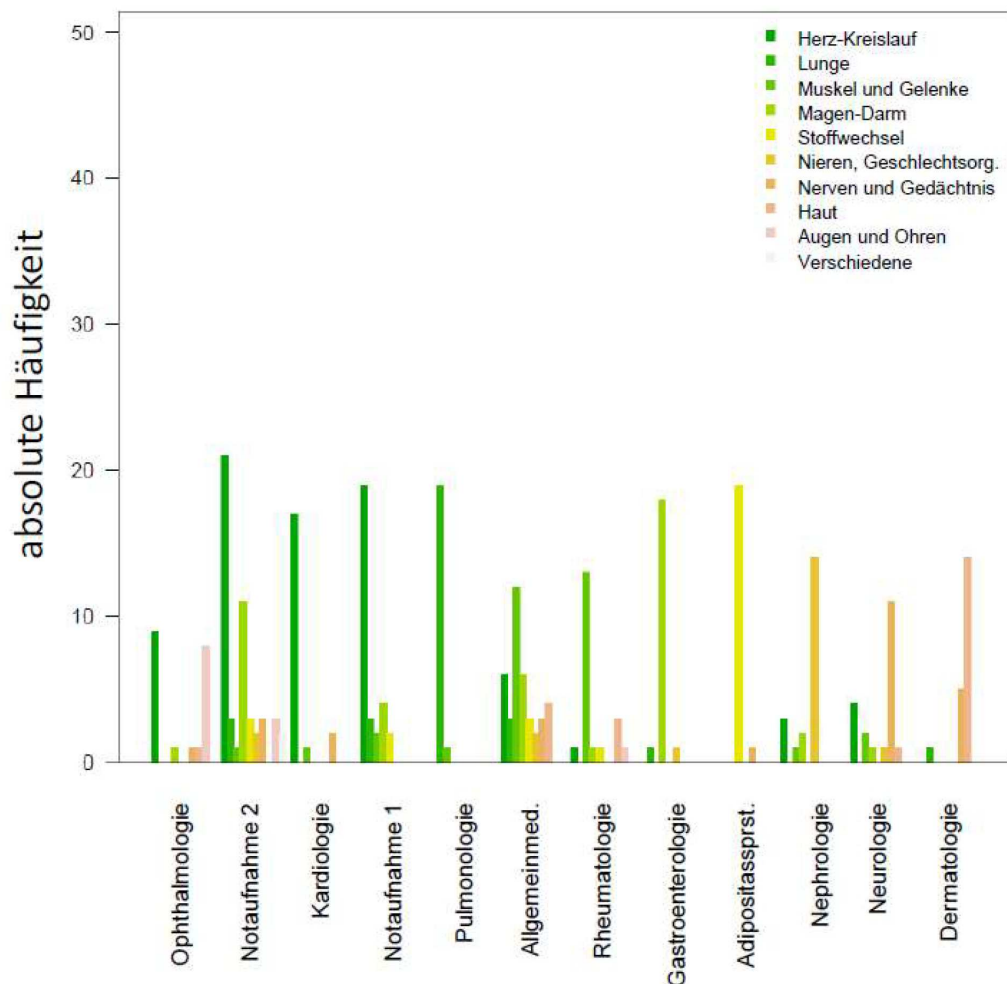


Abbildung 16: Verteilung der absoluten Häufigkeiten der OSSA Organdomänen mit den niedrigsten Gesundheitswerten in Abhängigkeit von der behandelnden Fachdisziplin; farblich kodiert

Hier wird das Behandlungsspektrum in den verschiedenen Instituten deutlich. In den Notaufnahmen und der allgemeinmedizinischen Praxis zeigt sich eine breite Fächerung der Organsystemdomänen. In den Notaufnahmen zeigt sich „Herz und Kreislauf“ als schlechtestes bewertetes Organsystem auf Patienten- und Arztseite (Notaufnahme 1: Patient: 67 %, Arzt 74,1 %; Notaufnahme 2: Patient 65,5 %, Arzt 71 %), in der Allgemeinmedizin stellt sich „Muskeln und Gelenke“ (Patient: 66,5 % Arzt: 80,2 %) als niedrigstes bewertetes Organsystem dar.

Entsprechend der spezialärztlichen Institute zeigt sich dort eine gehäufte Verteilung eines Organsystems.

3.4 Übereinstimmung der Einschätzung von Patienten und Ärzten bezüglich der Hauptdiagnosedomäne

Betrachten wir die Übereinstimmung der Hauptdiagnosedomäne in der Patienteneinschätzung mit der Einschätzung des Arztes. Diese misst im Median 74 %.

Tabelle 3: Übereinstimmung der Selbsteinschätzung mit der Hauptdiagnosedomäne

Klinik	Median
Rheuma 1	0,85
Lungenambulanz	0,7
Hautklinik	0,65
Nephrologie	0,6
Gastroenterologie	0,55
Notaufnahme 2	0,53
Notaufnahme 1	0,51
Kardiologie	0,5
Allgemeinmedizin	0,46
Augenklinik	0,45
Neurologie	0,4
Adipositas	0,15

Gesamt

0,74

Das heißt das vom Patienten gewählte Organsystem, also das subjektiv empfundene Hauptproblem hat eine hohe Übereinstimmung mit der vom Arzt gewählten Hauptdiagnosedomäne, bzw. dem entsprechenden Organsystem.

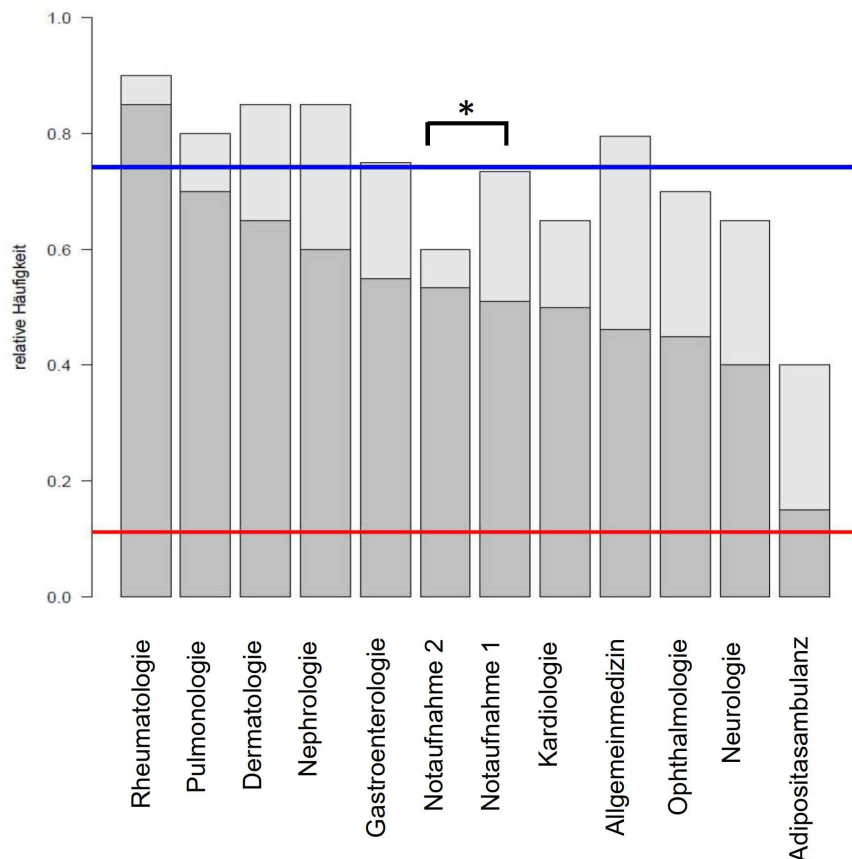


Abbildung 17: Listung der relativen Häufigkeiten in denen Arzt und Patient übereinstimmende Organdomänen als Domänen mit dem niedrigsten (assoziiert mit Hauptdiagnose) und zweit-niedrigsten Gesundheitszustand (assoziiert mit Nebendiagnose) identifiziert haben, gelistet nach den erhebenden Fachdisziplinen. Die blaue Linie bezeichnet eine relative Häufigkeit von 0,75. Die rote Linie bezeichnet die bei Zufallsverteilung, zu erwartende relative Häufigkeit von 0,11. Auffällig ist der signifikant Unterschied in der Arzt-Patienten Übereinstimmung hinsichtlich der Nebendiagnose assoziierten Organdomäne * $p < 0,01$.

Dieser hohe Übereinstimmung gilt jedoch in erster Linie der Hauptdiagnosedomäne. Die Übereinstimmung mit den Nebendiagnosedomänen fällt deutlich geringer aus (Abb. 17).

Im folgenden betrachten wir die Übereinstimmungen konkret am Beispiel der Notaufnahme. Gegenübergestellt werden Übereinstimmungen zwischen Patienteneinschätzung mit der Diagnose des Arztes, die Übereinstimmung der Patienteneinschätzung mit der Diagnose im Entlassungsbrief und die Übereinstimmung der Diagnose der ersten Arzteinschätzung mit der des späteren Entlassungsbriefes.

Tabelle 4: Übereinstimmung der Selbsteinschätzung mit der Hauptdiagnosedomäne (HD) und der Nebendiagnosedomäne (ND) in der Notaufnahme

Übereinstimmung in der Notaufnahme	Mittelwert
Übereinstimmung Patient und Arzt bei HD	64,7 %
Übereinstimmung Patient mit Arztbrief bei HD	55,8 %
Übereinstimmung erste Arzteinschätzung mit späterem Arztbrief bei HD	76,2 %
Übereinstimmung Patient und Arzt bei ND	19,6 %
Übereinstimmung Patient mit Arztbrief bei ND	21,5 %
Übereinstimmung erste Arzteinschätzung mit späterem Arztbriefe bei ND	24,4 %

Patient und Arzt haben in der Notaufnahme demnach eine relativ hohe Übereinstimmung in der Hauptdiagnosedomäne (64,7 %). Insbesondere gibt die hohe Übereinstimmung der Hauptdiagnosedomäne in der Arzteinschätzung mit der im Entlassungsbrief gelisteten Hauptdiagnose (76,2%) einen deutlichen Hinweis darauf, dass der Gesundheitszustand mit dem OSSA Fragebogen realistisch abgebildet werden

kann. Bei der Analyse der Nebendiagnosedomänen zeigt sich, dass die Übereinstimmungsraten deutlich geringer ausfallen.

3.5 Daten der Support Vector Machine (SVM)

Nun wird ein Modelle betrachtet, das basierend auf Patientendaten, die Arztentscheidung vorhersagt (Tabelle 5 und Abb. 18). Darüber hinaus wird ein Modell betrachtet, das basierend auf Arztdaten, die Einschätzung des Patienten vorhersagt (Tabelle 6 und Abb. 19).

Bei diesen ROC-Kurven betrachten wir die Fläche unter der Kurve (AUC) als Vorhersage-Güte. Zur Bewertung der Klassifikationsgüte wurde eine wiederholte Kreuzvalidierung (10-fach, 5 Wiederholungen) verwendet. Dies erlaubt eine Abschätzung der Verteilung des Gütemaßes.

Tabelle 5: AUCs geordnet nach Klinik, basierend auf Patientendaten, die die Arztentscheidungen vorhersagen

Pulmologie	AUC 0,67
Rheumatologie	AUC 0,58
Neurologie	AUC 0,54
Nephrologie	AUC 0,59
Augenklinik	AUC 0,76
Gastroenterologie	AUC 0,64
Dermatologie	AUC 0,73
Kardiologie	AUC 0,68

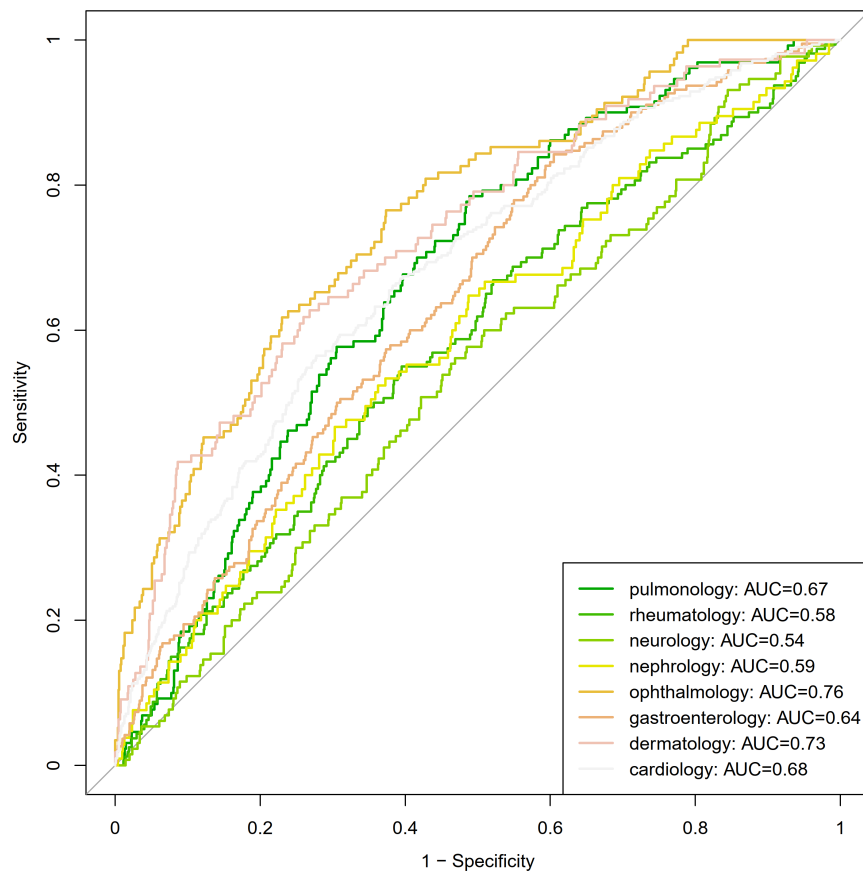


Abbildung 18: Darstellung der ROC-AUC Kurven bei Modellierung zur Prädiktion der Organdomäne mit der niedrigsten Gesundheits-einschätzung. Verwendet wurde das Instrument der Support Vector Machine (SVM, Repetitive Kreuzvalidierung, 10-Fach, 5-fach Wiederholung). Die Kurven sind nach Fachdisziplin farblich gekennzeichnet. Hier dienten die Patientenproben als Trainingsprobe.

Tabelle 6: AUCs geordnet nach Klinik, basierend auf Arztdaten, die die Patienteneinschätzung vorhersagen

Pulmologie	AUC 0,65
Rheumatologie	AUC 0,59
Neurologie	AUC 0,57
Nephrologie	AUC 0,57
Augenklinik	AUC 0,70
Gastroenterologie	AUC 0,66
Dermatologie	AUC 0,73
Kardiologie	AUC 0,66

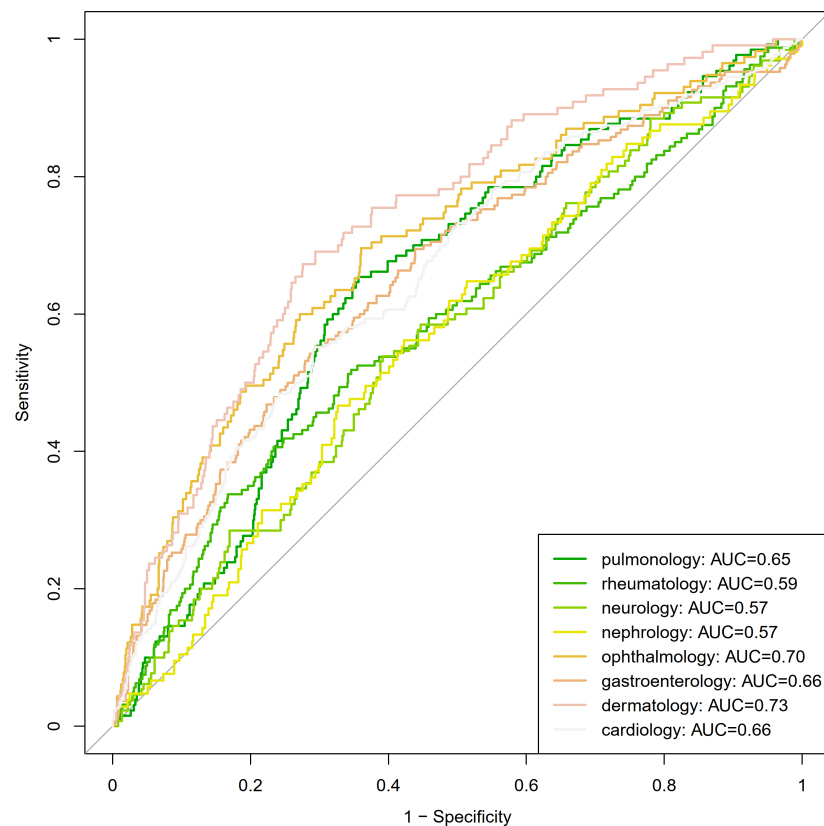


Abbildung 19: Darstellung der ROC-AUC Kurven bei Modellierung zur Prädiktion der Organdomäne mit der niedrigsten Gesundheitseinschätzung. Verwendet wurde das Instrument der Support Vector Machine (SVM, Repetitive Kreuzvalidierung, 10-Fach, 5-fach Wiederholung). Die Kurven sind nach Fachdisziplin farblich gekennzeichnet. Hier dienten die Arzteinschätzung als Trainingsprobe.

Der Fragebogen erlaubt eine korrekte Zuordnung des klinischen Problems des Patienten zu dem betroffenen Organsystem durch SVM-Analyse mit einer Mehrfach-AUC von bis zu 73,4 %.

Es lässt sich zusammenfassend bestätigen, dass dasjenige Organsystem, welches der Patient als niedrigstes bewertet hat, mit der ärztlich gestellten Hauptdiagnose assoziiert werden kann. Zudem kann die Nebendiagnose anhand der zweitniedrigsten Gesundheitseinschätzung abgebildet werden.

Dies bedeutet, dass mit dem OSSA-Fragebogen das von der Hauptdiagnose betroffene Organsystem korrekt erkannt werden kann

3.6 Zusammenfassung der Ergebnisse

Von den 410 OSSA- Fragebögen konnten 336 für die Berechnung der Support Vector Machine genutzt werden, da diese in kompletter Form vorlagen.

In der re-test-Variabilität bei 49 Patienten konnten wir keinen signifikanten Unterschied in der longitudinalen Befragung abbilden. Der Median der Erstmessung lag bei 85, in der Zweitmessung bei 87.

Das heißt sowohl die Patienten, als auch die Ärzte haben bei der zweiten Befragung, in der keine Änderung zu erwarten war, nahezu gleiche Einschätzungen zum Gesundheitszustand angegeben.

Mit einer Pearson Korrelation von $r=0,867$, $p<0,0001$ deutet der OSSA-Fragebogen auf eine gute Retest-Reliabilität hin, zudem besteht mit einem Cronbach alpha von 0,87 eine gute interne Konsistenz.

Der Fragebogen erlaubt eine korrekte Zuordnung des klinischen Problems des Patienten zu dem betroffenen Organsystem. Das Hauptproblem wird also gut abgebildet. Lediglich ein Unterschied zwischen Ausmaß der empfundenen Einschränkung zwischen Patient und Arzt liegt deutlich vor, denn die Patienten

schätzen ihren Gesundheitszustand in der Regel deutlich schlechter ein, als die Ärzte. Mit einem mittleren Prozentsatz von $66 \pm 1,17$ SE bzw. $74 \pm 0,98$ SE war die OSSA-Bewertung der Ärzte signifikant höher als die Bewertung der Patienten (p Wilcoxon $<0,001$).

Die Subjektivität der Patienten korreliert also gut mit der Objektivität des Arztes, wenn auch in unterschiedlicher Ausprägung.

Die Übereinstimmung zwischen Selbsteinschätzung des Patienten und Hauptdiagnosedomäne des Arztes liegt bei einem Median von 74.

Die ROC-AUC Kurven in der Support Vector Machine Untersuchung weisen Vorhersage-Wahrscheinlichkeiten zwischen 63,5 und 73,4 % auf. Dies bedeutet, dass mit dem OSSA-Fragebogen das von der Hauptdiagnosedomäne betroffene Organsystem korrekt erkannt werden kann.

Am Beispiel einer Notaufnahme mit einer Probangengröße von 42 Patienten konnten wir eine gute Übereinstimmung von 64,7% der Hauptdiagnosedomäne zwischen Patient u Arzt feststellen. Insbesondere gibt die hohe Übereinstimmung der Hauptdiagnosedomäne in der Arzteinschätzung mit der im Entlassungsbrief gelisteten Hauptdiagnose (76,2%) einen deutlichen Hinweis darauf, dass der Gesundheitszustand mit dem OSSA Fragebogen realistisch abgebildet werden kann. Bei der Analyse der Nebendiagnosedomänen zeigt sich, dass die Übereinstimmungsraten deutlich geringer ausfallen.

4 Diskussion

4.1 Diskussion der Methode

Patienten-reported-Outcomes (PRO) sind zunehmend wichtige Instrumente zur Unterstützung von diagnostischen und therapeutischen Entscheidungen geworden. Patienten-reported Outcomes werden dabei dadurch definiert, dass die dokumentierten Daten zum Gesundheitszustand als Patienteneinschätzung direkt und ohne Interpretation durch Ärzte oder andere dokumentiert werden (American Food and Drug Administration Guidance for Industry, 2009).

Ziel des Projektes war die Entwicklung eines interdisziplinär anwendbaren Patienten-Fragebogens, mit dem die Einschätzung des Gesundheitszustandes bezüglich verschiedener Organsysteme abgebildet werden kann. Diese Entwicklung ist notwendig, da eine einfache Verkettung von existierenden fachspezifischen Fragebögen im klinischen Alltag Patienten überfordert. Es wurde daher als Struktur in dem Fragebogen eine Verkettung von visuellen Analogskalen verwandt, um den Gesundheitszustand des Patienten bezüglich verschiedener Organsysteme abzubilden. Der Fragebogen wurde als Papierversion getestet und ausgewertet, ist aber konzipiert für eine digitale Anwendung und Auswertung.

In dieser Arbeit werden die ersten pilotierenden Untersuchungen zur Validierung und generellen Anwendbarkeit des Fragebogens vorgestellt.

In dieser Studie wurden zunächst 20 Patienten aus der Adipositasprechstunde und 29 Patienten aus der Rheumatologie longitudinal befragt in klinischen Umständen, die eine signifikante Veränderung des Gesundheitszustandes nicht erwarten ließen. Die Daten in Abbildung 1 zeigen eine hohe Retest-Reliabilität des OSSA-Fragebogens (Pearson Korrelation $r=0,867$, $p<0,0001$).

Weiter wurden Patienten aus verschiedenen medizinischen Disziplinen untersucht und parallel der OSSA-Fragebogen durch die behandelnden Ärzte ausgefüllt.

Zur Akzeptanz des Fragebogens wurden in dieser Studie keine strukturierten Interviews geführt. Aus Rücklauf- und Fehlerquote lassen sich jedoch Rückschlüsse darauf ziehen. Der Test ist auf einen kurzen zeitlichen Aufwand von ca. 2-3 Minuten konzipiert. Trotzdem war die Rücklaufquote auch bei Ärzten nicht vollständig. Diese Problematik zeigte sich vor allem in einer der Notaufnahmen. Dort wurden 98 % der Patientenfragebögen ausgefüllt, jedoch nur 30 % der Arztfragebögen. In der anderen Notaufnahme sahen die Zahlen jedoch deutlich anders aus. Hier haben 93 % der Patienten und 98 % der Ärzte die Fragebögen ausgefüllt. Die meisten Rücklaufzahlen waren jedoch durchaus akzeptabel, weshalb es sich in der einen Notaufnahme nur fraglich um eine repräsentative Quote handelt und am ehesten um eine Ausnahme handelt. Diese Situation zeigt jedoch wie zeitkritisch die Akzeptanz eines Fragebogens ist.

Insgesamt waren 336 (82 %) der ausgeteilten Fragebögen in allen Belangen vollständig vom Patienten und Arzt ausgefüllt und für die Endauswertung mittels Support Vector Machine verfügbar.

Inhaltliche oder Verständnisprobleme scheint der Test nur wenige zu bereiten. Zum einen wurde gelegentlich vergessen die zweite Seite des Fragebogens auszufüllen. Ein paar wenige Ausnahmen interpretierten die VA-Skala trotz Beschriftung falsch, sodass 0 % als volle Gesundheit gewertet wurde und diese Fragebögen nicht ausgewertet werden konnten.

Beiden Problemen wird in der nachfolgend entwickelten digitalen Tablett basierten Form des OSSA Rechnung getragen. Ein Weiterspringen bei unbeantworteten Fragen wird programmseits verhindert. Zusätzlich wurde der Schiebebalken der visuellen Analogskalen mit einem farblichen Verlauf gekennzeichnet (Grün= guter Gesundheitszustand; rot= schlechter Gesundheitszustand), so dass ein Vertauschen der Skalenrichtung vermieden wird.

Es wurden bewusst nur die drei Orientierungshilfen „75 % Behandlungsbedarf, 50 % dringlicher Behandlungsbedarf, 25 % Notfall“ auf der Skala verankert, damit der Patient intuitiver antworten konnte. So wurde vermieden genauere Vorgaben zu machen, wodurch die Subjektivität bewahrt werden konnte. Durch diese Standardisierung der Fragebögen kann eine hohe Durchführungsobjektivität und Auswertungsobjektivität erzielt werden.

Als Anhalt für eine gute interne Konsistenz des OSSA lag die aus den vollständig vorliegenden Daten ermittelte Cronbach's alpha Maßzahl bei 0,87 (Nunnally, J. C., 1978).

Systemimmanent können mit dem OSSA Fragebogen die gesundheitlichen Einschränkungen bestimmter Erkrankungen, beispielsweise aus dem Bereich der Hämatologie, weniger gut dargestellt werden. Bei Vorliegen einer Erkrankung, die aus Sicht des Patienten nicht einem der aufgezählten Organsystemen zugeordnet werden konnte, kann diese Erkrankung nur im „allgemeinen Gesundheitszustand“ abgebildet werden.

Die aufgeführten Testcharakteristika des OSSA Fragebogens weisen auf eine gute Retest-Reliabilität und gute interne Konsistenz hin. Es bleiben jedoch noch eine Reihe von Fragen offen, deren Beantwortung für eine vollständige Validierung notwendig ist.

Das Alter der Probanden wurde zum derzeitigen Stand des Projektes nicht berücksichtigt. Die Zusammenhänge zwischen Einschätzung des Gesundheitszustandes und Alter sind klar belegt

„The discrepancy found in ratings of health between physicians and the subjects themselves can be explained by different appreciation of ageing in the assessment of health status. After adjustment for medical conditions, age did not relate to self-perceived health, whereas there was a significant association between age and doctors' ratings.“ (Kivinen, P., et al., 1998).

Weiter bestehen Abhängigkeiten der Gesundheitseinschätzung von Geschlecht (Sieverding, M., 1998), sozioökonomischem Status und Bildung (Subramanian, 2010).

Die Untersuchung der Abhängigkeit des OSSA-Fragebogens von Alter und Geschlecht sowie soziodemographischen Faktoren und eine Korrelation mit dem bereits beschriebenen EQ-5D Fragebogen hätte jedoch den Rahmen des vorgelegten Pilotprojektes weit überschritten. Diese Untersuchungen sollen Gegenstand von Folgeprojekten sein.

Ebenso wurden Patienten mit psychischen Erkrankungen in der vorgelegten als Pilotprojekt konzipierten Arbeit nicht untersucht. Bei diesen Erkrankungen erscheint die generalisierte Anwendbarkeit des Fragebogens zweifelhaft. Die Prüfung des Fragebogens in Zusammenhang mit einem Depressionsfragebogen bei Patienten mit Depressionen könnte ebenfalls eins der zukünftigen Folgeprojekte darstellen.

Die wichtigste Frage für Folgeprojekte ist jedoch die nach der longitudinalen Diskriminationsfähigkeit des OSSA bei Wiederholungsmessungen mit verändertem Gesundheitszustand, denn die Befragten sind gut in der Lage, Veränderungen ihrer Gesundheit im zeitlichen Verlauf wahrzunehmen (Carstensen, 2016).

4.2 Diskussion der Ergebnisse

Bei der Analyse der Patienteneinschätzung und der Arzteinschätzung besteht eine hohe Korrelation zwischen Patienteneinschätzung und Arzteinschätzung (Median 74). In verschiedenen Fachdisziplinen fällt auf, die Einschätzung der Ärzte, bis auf wenige Ausnahmen, etwas positiver ausfällt als die der Patienten. Bei der Analyse der Einschätzung des allgemeinen Gesundheitszustandes ergibt sich ein signifikanter Unterschied (Arzteinschätzung $74 \% \pm 0,98 \text{ SEM}$ vs. Patienteneinschätzung $66 \% \pm 1,17 \text{ SEM}$, $p < 0,001$). Derartige Unterschiede werden auch bei der Auswertung von Registerdaten berichtet (Gvozdenovic E. et al., 2014).

Diese Beobachtung wird lediglich von zwei Ausnahmen durchbrochen. In der Nephrologie behandelte Patienten schätzten den Gesundheitszustand bezüglich Nieren- und Geschlechtsorganen günstiger ein als die behandelnden Ärzte (Abb. 9). Dies ist jedoch eine plausible Beobachtung, da eine Unterschätzung von Nierenerkrankungen durch die Betroffenen aufgrund der Schmerzlosigkeit hinlänglich bekannt ist.

Eine weitere Ausnahme stellen Patienten aus der Adipositasprechstunde dar. Hier wird der Gesundheitszustand bezüglich des Stoffwechsels von den Ärzten ungünstiger beurteilt (Abb. 13). Der muskuloskelettale Gesundheitszustand wird von dieser Patientengruppe demgegenüber schlechter beurteilt als das passende Organsystem „Stoffwechsel“. Auch diese Beobachtung ist hochplausibel und spiegelt die klinische Situation realistisch wieder.

Die Daten aus den verschiedenen klinischen Fachdisziplinen bilden gut die zu erwartenden Organsysteme ab. Bei der Testauswertung des OSSA ist zu erwarten, dass die Hauptdiagnosen zu den niedrigsten Einschätzungswerten in der entsprechenden Organsystemdomäne des OSSA Fragebogens führt. Diese Erwartung ist in den Abbildungen 4 bis 15 gut nachzuvollziehen, in denen die Mittelwerte \pm SEM der verschiedenen Organdomänen des Fragebogens in Abhängigkeit der klinischen Fachdisziplin gezeigt werden. Abbildung 16 zeigt die Häufigkeitsverteilung der „schlechtesten“ Gesundheitseinschätzungen in den verschiedenen Fachdisziplinen. Erneut zeigt sich eine realistische Abbildung des Behandlungsspektrums in den verschiedenen Disziplinen. So ist besteht in den Notaufnahmen und in der Allgemeinmedizinischen Praxis eine breitere Fächerung der entsprechenden Organsystemdomänen, während in den spezialisierteren Institutionen eine schmale Häufigkeitsverteilung mit Häufung der der Institution zuzuordnenden Organdomäne festzustellen ist. Die beiden Notaufnahmen zeigen dabei ein ähnliches Verteilungsspektrum (Abb. 11 und Abb. 12). In beiden Notaufnahmen weist das Organsystem „Herz und Kreislauf“ sowohl beim Arzt, als auch beim Patient jeweils die niedrigste Prozentzahl auf (Notaufnahme 1: Patient: 67 %, Arzt 74,1 %; Notaufnahme 2: Patient 65,5 %, Arzt 71 %). Dies entspricht den aktuellen allgemeinen Beobachtungen. Denn der Myokardinfarkt ist in

Deutschland mit Abstand die häufigste Todesursache (Post, F., et al, 2007).

Zudem beträgt der Anteil der Patientengruppen, die an Thoraxschmerzen leiden mehr als 20 % in der Notaufnahme (Michalski, T., et al, 2012).

In der Allgemeinmedizinischen Praxis stellt sich Muskeln und Gelenke als schlecht bewertetes Organsystem dar, sowohl in der Einschätzung des Patienten (66,57 %), als bei der Arzteinschätzung (80,25 %). Laut dem deutschen Ärzteblatt (Oktober 2016) zeigt sich entsprechend, dass Muskel-, Skelett- und Bindegewebserkrankungen die häufigsten Diagnosen in deutschen Arztpraxen sind, was diese Beobachtung bestätigen kann.

In einem weiteren Schritt wurden untersucht, ob mittels einem linearen Modell einer Support Vector Machine (SVM) die Daten geeignet sind die O SSA- Organdomäne der Hauptdiagnose des Patienten korrekt zu erkennen. Die Ergebnisse sind in Abbildung 18 und 19 dokumentiert. Die ROC-AUC Kurven weisen Vorhersage-Wahrscheinlichkeiten zwischen 63,5 und 73,4 % auf. Dies bedeutet, dass mit dem O SSA-Fragebogen das von der Hauptdiagnose betroffene Organsystem korrekt erkannt werden kann. Zu bedenken ist dabei die für derartige Berechnung relative kleine Anzahl von 336 untersuchten Patienten. Diese Daten bestätigen, die vielversprechenden Testcharakteristika des O SSA-Fragebogens in dieser Pilotstudie.

Bei einer Teilstichprobe von 42 Patienten aus Einer der Notaufnahmen wurde die in der Ersteinschätzung des Arztes im O SSA ermittelte Organdomäne mit der Hauptdiagnose des Entlassungsbriefes verglichen (Tabelle 4). In 32 von 42 (76,2%) bestand eine Übereinstimmung mit der der Hauptdiagnose zuzuordnenden Organdomäne. Diese Übereinstimmung erscheint im Setting einer Notaufnahme realistisch.

Ähnlich der Auswertung der Organdomäne der Hauptdiagnose können in der Auswertung auch Nebendiagnosen abgebildet werden, indem eine Nebendiagnose in der Organdomäne mit der zweitniedrigsten Gesundheitseinschätzung vermutet wird.

Wenn die so ermittelten jeweiligen Organdomänen des Tests bei Einschätzung bei Patient und Arzt übereinstimmen, kann von einer „richtigen“ Erkenntnis der Organdomäne bezüglich einer Haupt- und einer Nebendiagnose ausgegangen werden.

In Abbildung 17 sind die relativen Häufigkeiten der Arzt- und Patientenübereinstimmungen für die im OSSA als Haupt- und Nebendiagnose ermittelten Organdomänen dargestellt.

Es zeigt sich, dass die Übereinstimmungsraten für die Nebendiagnosedomänen deutlich geringer sind als für die Hauptdiagnosedomäne. Die Übereinstimmungshäufigkeiten liegen aber deutlich über der Zufallswahrscheinlichkeit von 0,11. Die Unterschiede in den einzelnen Fachdisziplinen deuten auf Unterschiede Krankheitswahrnehmung hin.

Auffällig ist bei dem Vergleich der Daten der beiden Notaufnahmen, dass die Häufigkeit der Arzt-/Patientenübereinstimmung bezüglich der Nebendiagnosedomäne sich signifikant unterscheidet (Wicoxon $p < 0,001$) bei gleichen Übereinstimmungsraten bezüglich der Hauptdiagnosedomäne. Die Ursache hierfür muss derzeit unklar bleiben.

4.3 Zusammenfassung der Diskussion und Ausblick

Der neu entwickelte OSSA-Fragebogen zeigt in den ersten in dieser Arbeit vorgelegten Validierungsuntersuchungen ein vielversprechende Testcharakteristika hinsichtlich Retest-Reliabilität und interner Konsistenz.

Durch die Standardisierung des Fragebogens unter Verwendung von „verankerten“ visuellen Analogskalen ermöglicht eine hohe Durchführungsobjektivität und Auswertungsobjektivität.

Anhand des Fragebogens können zielführend dargestellt werden, dass der Patient und sein behandelnder Arzt eine hohe Übereinstimmung in der Auffindung der Organdomänen der Hauptdiagnose haben. Auch wenn sich dabei die Patienten fast

regelmäßig schlechter einschätzten als die Ärzte, wird das gleiche Problem abgebildet. Das heißt die Subjektivität der Patienten und die Objektivität der Ärzte korrelieren gut.

Bei Einsatz des Verfahrens der Support Vector Machine auf Basis von Datensätzen von 336 Patienten lagen die ermittelten ROC-AUC Werte bei 63,5 bis 73,4 % für die Prädiktion der Organdomäne der Hauptdiagnose. Dies deutet auf eine Auswertungsmöglichkeit mittels selbstlernender Systeme hin.

Der OSSA-Fragebogen kann durch zusätzlich Hinweise auf Gesundheitsprobleme in verschiedenen Organdomänen behandelnden Ärzte wichtige Zusatzinformationen liefern und so z.B. weitergehende Anamnese und ggf. diagnostische Maßnahmen triggern. Eine weitere Anwendung des OSSA Fragebogens wäre als Qualitätssicherungsinstrument denkbar, bei paralleler Anwendung durch Ärzte und Patienten.

Für die weitere Validierung sind noch viele Fragen zu klären. Die Abhängigkeit der im OSSA-Fragebogen ermittelten Aussagen von Alter, Geschlecht, psychischer Verfassung und sozioökonomischem Status steht aus. Ebenso muss die Diskriminationsfähigkeit bei longitudinaler Anwendung geprüft werden. Die vorliegenden Daten rechtfertigen jedoch weitere Validierungsschritte in Folgestudien.

Zwischenzeitlich ist der OSSA-Fragebogen als Tablett-basierte Version verfügbar, was Anwendung und Auswertung weiter erleichtert.

5 Zusammenfassung

Bisher gibt es kein standardisiertes Verfahren zur Darstellung des subjektiven Gesundheitszustand der Patienten bezogen auf spezifische Organsysteme.

Die wachsende Multimorbidität in Zeiten des demographischen Wandels stellt eine große Herausforderung dar.

Ziel ist es, zukünftig mittels Fragebogen den behandelnden Arzt in Entscheidung und Resourcenallokation sinnvoll zu unterstützen.

In dieser Arbeit wurde daher ein standardisierter Fragebogen entwickelt, der mittels visueller Analogskala valide den subjektiven Gesundheitszustand abbilden kann. Die zehn Fragen beziehen sich dabei auf einzelne Organsysteme (Organ Specific Self-Assesment).

Zur erweiterten Testung und Evaluation werden neben den Patienten die behandelnden Ärzte mit dem gleichen Fragebogen befragt. So wird die Subjektivität der Patienten der Objektivität der Ärzte gegenübergestellt und verglichen. Zudem geben die Ärzte die fünf wichtigsten Diagnosen in der Reihenfolge ihrer Bedeutung für den Gesundheitszustand des Patienten an.

Die Studienpopulation betrug 410 Patienten plus deren behandelnde Ärzte.

Der Test zeigt eine gute Reproduzierbarkeit. Mit einem mittleren Prozentsatz von $66 \pm 1,17$ SE bzw. $74 \pm 0,98$ SE ist die OSSA-Bewertung der Ärzte signifikant höher als die Bewertung der Patienten (p Wilcoxon $<0,001$). Modelle, die Hauptdiagnosen vorhersagen, werden so konstruiert und geschätzt, dass sie mit Multi-Klassen-AUCs von 63,5 % bzw. 73,4 % (basierend auf der OSSA des Patienten bzw. des Arztes) funktionieren.

Schlussfolgerung:

In diesem Vorversuch mit geringer Probengröße zeigt der Fragebogen POPGEN-OSSA eine gute Reproduzierbarkeit und ermöglicht eine korrekte Zuordnung des klinischen Problems des Patienten zum beteiligten Organsystem durch SVM-Analyse mit einer AUC in mehreren Klassen von bis zu 73,4 %.

Der OSSA-Fragebogen zeigt in den ersten in dieser Arbeit vorgelegten Validierungsuntersuchungen ein vielversprechende Testcharakteristika hinsichtlich Retest-Reliabilität und interner Konsistenz

Diese Daten verdienen eine weitere Untersuchung an einer größeren Patientenkohorte. Eine APP-basierte Version wird bereits entwickelt.

6 Literaturverzeichnis

- Altenburg, H-P.: "Analyse von ROC-Kurven mit Hilfe von SAS." (2009)
- American Food and Drug Administration Guidance for Industry: Patient-Reported Outcome Measures: Use in Medical Product Development to Support Labeling Claims 2009, <https://www.fda.gov/media/77832/download>, (Aufruf am 27.08.2019)
- Boge, K-S.: Evaluation der Patiententoleranz von Wachoperationen zerebraler Gliome in Bezug auf das Angstempfinden und die Belastungssituation. Diss. Universitäts-und Landesbibliothek der Heinrich-Heine-Universität Düsseldorf, 2015
- Borg, I.: "Multidimensionale Skalierung." Handbuch der sozialwissenschaftlichen Datenanalyse. VS Verlag für Sozialwissenschaften, 2010: 391-418
- Broderick, J.E., et al: "Recalled pain ratings: a complex and poorly defined task." The Journal of Pain 7.2 (2006): 142-149
- Clark, P., et al: "Learning from pain scales: patient perspective." The Journal of Rheumatology 30.7 (2003): 1584-1588
- Carstensen, J.: "Die Messung von Gesundheit." Handbuch Gesundheitssoziologie 2016: 1-21
- Cortes, C., and Vladimir V.: "Support-vector networks." Machine learning 20.3 (1995): 273-297
- deutsche Ärzteblatt 10/16
- Eccleston, C., et al: "Attention and somatic awareness in chronic pain." Pain 72.1-2 (1997): 209-215

- Ferrell B.A.: Pain management. Clin Geriatr Med 2000;16(4):853-874
- Fischer, J.: Support Vector Machines (SVM). In Seminar" Statistische Lerntheorie und ihre Anwendungen, 2007
- Funke, Frederik. "Internet-based measurement with visual analogue scales: An experimental investigation." Online im Internet: <http://nbn-resolving.de/urn:nbn:de:bsz> (2010)
- Gagliese, L., et Melzack, R.: "Chronic pain in elderly people." Pain 70.1, 1997: 3-14
- Gvozdenović, E., et al. "Assessment of global disease activity in RA patients monitored in the METEOR database: the patient's versus the rheumatologist's opinion." Clinical rheumatology 33.4 (2014): 461-466
- Hand, D. J., Till, R. J.: "A simple generalisation of the area under the ROC curve for multiple class classification problems." Machine learning 45.2 (2001): 171-186
- Hayes, M. H.: "Experimental developement of the graphics rating method." Physiol Bull 18 (1921): 98-99
- Heller, Gillian Z., Maurizio Manuguerra, and Roberta Chow. "How to analyze the Visual Analogue Scale: Myths, truths and clinical relevance." Scandinavian journal of pain 13.1 (2016): 67-75
- Helmert, U.: "Subjektive Einschätzung der Gesundheit und Mortalitätsentwicklung." Das Gesundheitswesen 65.01 (2003): 47-54
- Herr, Keela A., and Linda Garand. "Assessment and measurement of pain in older adults." Clinics in geriatric medicine 17.3 (2001): 457-478
- Huskisson, E.C.: "Measurement of pain." The lancet 304.7889 (1974): 1127-1131

- Karner, J.: Die Abbildung chronischer Schmerzen anhand von validierten Fragebögen: eine qualitative Studie bei älteren Patienten mit chronischen Schmerzen der Halswirbelsäule.“ Diss. 2012
- Katz, J., Melzack, R.: "Measurement of pain." *Surgical Clinics of North America* 79.2 (1999): 231-252
- Kivinen, P., et al.: "Self-rated health, physician-rated health and associated factors among elderly men: the Finnish cohorts of the Seven Countries Study." *Age and ageing* 27.1 (1998): 41-47
- König, Hans Helmut, Sebastian Bernert, and Matthias C. Angermeyer. "Gesundheitszustand der deutschen Bevölkerung: Ergebnisse einer repräsentativen Befragung mit dem EuroQol-Instrument." *Das Gesundheitswesen* 67.03 (2005): 173-182
- Leshem-Rubinow, E., et al.: "Self-rated health is associated with elevated C-reactive protein even among apparently healthy individuals." *Isr Med Assoc J* 17.4 (2015): 213-218
- Litcher-Kelly, Leighann, et al. "A systematic review of measures used to assess chronic musculoskeletal pain in clinical and randomized controlled clinical trials." *The Journal of Pain* 8.12 (2007): 906-913
- Marengoni, Alessandra, et al. "Aging with multimorbidity: a systematic review of the literature." *Ageing research reviews* 10.4 (2011): 430-439
- Meder, M., and E. Farin. "Akzeptanz und Verständlichkeit verschiedener Methoden der Gesundheitsbewertung bei chronisch Kranken: Willingness to pay, visuelle Analogskala und verbale Ratingskala." *Das Gesundheitswesen* 71.11 (2009): e62-e71

- Meyer, T., R. Deck, and H. Raspe. "Validity of patient self report data in rehabilitation research: identifying circumstances of the patients' completion of questionnaires." *Die Rehabilitation* 45.2 (2006): 118-127
- Miaskowski C., et al: *Principles of Analgesic Use in the Treatment of Acute Pain and Cancer Pain*. American Pain Society, 2008
- Michalski, T., et al. "Kardiologie." *Journal of Cardiology* 19 (2012): 11-12
- Morfeld, M., and M. Bullinger: "Der SF-36 Health Survey zur Erhebung und Dokumentation gesundheitsbezogener Lebensqualität." *Physikalische Medizin, Rehabilitationsmedizin, Kurortmedizin* 18.05 (2008): 250-255
- Nietzsche, Friedrich: *Die fröhliche Wissenschaft* KSA 1882 ergänzt 1887
- Nunnally, J. C. "Psychometric theory (2nd edit.) mcgraw-hill." Hillsdale, NJ 416 (1978)
- Post, Felix, Sabine Genth-Zotz, and T. Munzel. "Aktueller Stellenwert einer Chest Pain Unit in Deutschland." *Herz* 32.5 (2007): 435-437
- Robert Koch-Institut: *Public Use File BGS98, Datensatz des Bayerischen Gesundheitssurvey 1998*
- Sieverding, Monika. "Sind Frauen weniger gesund als Männer? Überprüfung einer verbreiteten Annahme anhand neuerer Befunde." *Kölner Zeitschrift für Soziologie und Sozialpsychologie* 50.3 (1998): 471-489
- Subramanian, S. V., et al: "Self-reported health assessments in the 2002 World Health Survey: how do they correlate with education?." *Bulletin of the World Health Organization* 88 (2010): 131-138
- Stinson, J. N., et al: "Systematic review of the psychometric properties, interpretability and feasibility of self-report pain intensity measures for use in clinical trials in children and adolescents." *Pain* 125.1-2 (2006): 143-157

- Wasilewski, D., et al: "Depression and pain: Polish epidemiologic study." *Psychiatria polska* 44.3 (2010): 435-445
- Weltgesundheitsorganisation: Verfassung der Weltgesundheitsorganisation: Unterzeichnet in New York am 22. Juli 1946, Ratifikationsurkunde von der Schweiz hinterlegt am 29. März 1947, von der Bundesversammlung genehmigt am 19. Dezember 1946, für die Schweiz in Kraft getreten am 7. April 1948
- Wiitavaara, B., et al: "How well do questionnaires on symptoms in neck-shoulder disorders capture the experiences of those who suffer from neck-shoulder disorders? A content analysis of questionnaires and interviews." *BMC musculoskeletal disorders* 10.1 (2009): 30

7 Danksagung

Zuerst möchte ich mich bei meinem Betreuer PD Dr. Rainald Zeuner für sein persönliches Engagement und seine vielfältige Unterstützung ganz herzlich bedanken. Er hat mir diese Doktorarbeit ermöglicht und mich intensiv und stets zuverlässig betreut. Dem Direktor der Medizinischen Klinik I Prof. Dr. med. S. Schreiber, sowie der Leitung der Sektion Rheumatologie Prof. Dr. med. Hoyer und der ehemaligen Leiter Prof. Dr. med. Schröder, gilt selbstverständlich ebenfalls ein großer Dank.

Dr. rer. nat. Matthias Hübenthal (Department of Dermatology, Venereology and Allergy University Hospital Schleswig-Holstein, Campus Kiel) hat mich im Bereich Statistik unterstützt, wofür ich mich besonders bedanke.

Ein weiterer Dank geht an die verschiedenen Institute des UKSH Kiel und deren Leiter, die mich bei der Probandenrekrutierung unterstützt haben: Prof. rer. nat. A. Franke (Institut für klinische Molekularbiologie), Prof. Dr. med. M. Laudes (Klinik für Innere Medizin I, Endokrinologie, Diabetologie, klinische Ernährungsmedizin), PD. Dr. Nikolaus (Klinik für Innere Medizin I, Gastroenterologie), Prof. Dr. B. Bewig (Klinik für Innere Medizin I, Pneumologie), Prof. Dr. Frey (Klinik für Innere Medizin III, Kardiologie), Prof. Dr. Berg (Klinik für Neurologie, Uni Kiel), Prof. Dr. Kunzendorf (Klinik für Innere Medizin IV, Nephrologie), Prof. Dr. Schwarz (Klinik der Dermatologie, Uni Kiel), Prof. Dr. Roider (Augenklinik, Uni Kiel), H. Herzog (internistische Aufnahmestation des städtischen Krankenhauses Kiel), Dr. med. Lohmann aus der allgemeinmedizinischen Praxis in Preetz.

Den zahlreichen Ärzten und Patienten, welche die Fragebögen ausgefüllt haben, gilt selbstverständlich ein ganz besonderer Dank. Ohne die Probanden und die Unterstützung derer Ärzte wäre diese Arbeit nicht möglich gewesen.

Meine Eltern haben mich stets unterstützt und mir die Ausbildung ermöglicht, wofür ich mich ganz herzlich bedanke.

8 Anhang

Pat ID:

POPGEN-OSSA Arztbogen

1. Wie beurteilen sie den allgemeinen Gesundheitszustand?

Interventionsbedarf 75% Dringlicher Interventionsbedarf 50% Notfall 25%



2. Wie beurteilen sie den Gesundheitszustand von Herz- und Kreislauf?

Interventionsbedarf 75% Dringlicher Interventionsbedarf 50% Notfall 25%
Beispiele: NYHA 3: 30%, stabil eingestellte arterielle Hypertonie ohne Komplikationen: 90%,
Belastungsdyspnoe beginnend: 75%



3. Wie beurteilen sie den Gesundheitszustand der Lungen?

Interventionsbedarf 75% Dringlicher Interventionsbedarf 50% Notfall 25%
Beispiele: COPD GOLD 3: 20%, Belastungsdyspnoe beginnend: 75%, 02-Hemtherapie: 10%



4. Wie beurteilen sie den Gesundheitszustand von Muskeln und Gelenken?

Interventionsbedarf 75% Dringlicher Interventionsbedarf 50% Notfall 25%
Beispiel: dokumentierbare Arthritissymptome (Druckschmerz und Schwellung)
von einem Gelenk als Erstmanifestation: 60%



5. Wie beurteilen sie ihren Gesundheitszustand von Magen, Darm und Verdauungssystem?

Interventionsbedarf 75% Dringlicher Interventionsbedarf 50% Notfall 25%
Beispiele: klinisch signifikanter unbeschränkter Gewichtsverlust: 75%,
Diarrhoeen mit beginnenden klinischen Exsikkationssymptomen: 50%



6. Wie beurteilen sie den Gesundheitszustand des Stoffwechsels?

Interventionsbedarf 75% Dringlicher Interventionsbedarf 50% Notfall 25%
Beispiel: Diabetes mit chronischen Komplikationen (z.B. Fußsyndrom, Retinopathie): 50%



7. Wie beurteilen sie den Gesundheitszustand von Nieren und Geschlechtsorganen?

Gesund



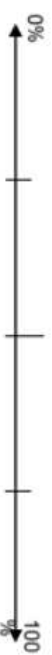
Interventionsbedarf 75% Dringlicher Interventionsbedarf 50% Notfall 25%
 Beispiele: Dialysepflichtigkeit: 100%, symptomatischer Harnwegsinfekt mit Fieber: 70%,
 chronische Anämie bei Menorrhagie: 70%,

8. Wie beurteilen sie den Gesundheitszustand von Nerven und Gedächtnis?



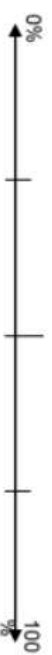
Interventionsbedarf 75% Dringlicher Interventionsbedarf 50% Notfall 25%
 Beispiele: Polyneuropathie mit beginnender Gehbehinderung: 70%,
 Chronische Hämiparese mit Gehbehinderung 30%, ohne Gehbehinderung 50%

9. Wie beurteilen sie den Gesundheitszustand der Haut?



Interventionsbedarf 75% Dringlicher Interventionsbedarf 50% Notfall 25%
 Beispiele: Hautläsionen an sichtbaren Körperstellen die im Alltag auffallen: 70%

10. Wie beurteilen sie den Gesundheitszustand von Auge und Gehör?



Interventionsbedarf 75% Dringlicher Interventionsbedarf 50% Notfall 25%
 Beispiele: Brillenträger ohne besonders starke Korrektur: 85%,
 Schwerhörigkeit mit dauerhaftem Bedarf eines Hörgerätes: 50%

Nennen Sie die fünf „wichtigsten“ Diagnosen nach Bedeutung für den Patienten

- 1 _____
- 2 _____
- 3 _____
- 4 _____
- 5 _____

Pat ID:

POPGEN-OSSA Fragebogen

Gesund

1. Wie empfinden sie den allgemeinen Gesundheitszustand?

Behandlungsbedarf 75% Dringlicher Behandlungsbedarf 50% Notfall 25%



2. Wie empfinden sie den Gesundheitszustand von Herz- und Kreislauf?

Behandlungsbedarf 75% Dringlicher Behandlungsbedarf 50% Notfall 25%



3. Wie empfinden sie den Gesundheitszustand der Lungen?

Behandlungsbedarf 75% Dringlicher Behandlungsbedarf 50% Notfall 25%



4. Wie empfinden sie den Gesundheitszustand von Muskeln und Gelenken ?

Behandlungsbedarf 75% Dringlicher Behandlungsbedarf 50% Notfall 25%



5. Wie empfinden sie ihren Gesundheitszustand von Magen, Darm und Verdauungssystem?

Behandlungsbedarf 75% Dringlicher Behandlungsbedarf 50% Notfall 25%



6. Wie empfinden sie den Gesundheitszustand des Stoffwechsels?

Behandlungsbedarf 75% Dringlicher Behandlungsbedarf 50% Notfall 25%



7. Wie empfinden sie den Gesundheitszustand von Nieren und Geschlechtsorganen?

Behandlungsbedarf 75% Dringlicher Behandlungsbedarf 50% Notfall 25%



8. Wie empfinden sie den Gesundheitszustand von Nerven und Gedächtnis?

Behandlungsbedarf 75% Dringlicher Behandlungsbedarf 50% Notfall 25%



9. Wie empfinden sie den Gesundheitszustand der Haut?

Behandlungsbedarf 75% Dringlicher Behandlungsbedarf 50% Notfall 25%



10. Wie empfinden sie den Gesundheitszustand von Auge und Gehör?

Behandlungsbedarf 75% Dringlicher Behandlungsbedarf 50% Notfall 25%





POGEN-OSSA: DEVELOPMENT OF AN ORGAN SPECIFIC SELF ASSESSMENT (OSSA®*) FOR INTERDISCIPLINARY DOCUMENTATION OF PATIENT REPORTED CLINICAL OUTCOMES

THU0682

R. Zeuner^{1,*}, U. Gsell², M. Hübenthal³, S. Schreiber⁴, A. Franke³, M. Laudes⁴, J. O. Schröder¹

¹Section of Rheumatology, Clinic for Internal Medicine I, University Hospital Schleswig-Holstein, Campus Kiel, Kiel, 24Klinik Eichstädt, Eichstätt, Germany
²Clinic for Internal Medicine I, University Hospital Schleswig-Holstein, Campus Kiel, Kiel, Germany



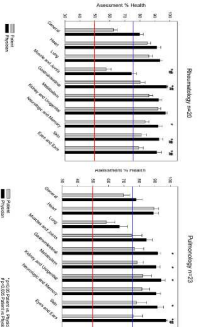
Background: Patient reported outcome measures are comprised of either sets of questionnaires or patient global assessment based on visual analogue scale (VAS). These patient-reported outcome measures lack accuracy and/or clinical feasibility when comparing heterogeneous patient groups with different diseases, or when characterizing patients with systemic disease involving different organ systems. Therefore we developed a clinical feasible patient-reported outcome measure that is based on VAS assessment of different organ systems.

Methods: The questionnaire consists of 10 items – one concerning the general health and nine addressing the health of different organ systems.

- How do you rate your general health?
- How do you rate the health of your heart and vessels?
- How do you rate the health of your muscles and joints?
- How do you rate the health of your lungs?
- How do you rate the health of your gastrointestinal system?
- How do you rate the health of your kidneys and urogenital system?
- How do you rate the health of your neural system and memory?
- How do you rate the health of your skin?
- How do you rate the health of your ears and eyes?

Patients were asked to rate each item, using a VAS-scale from the 'right = 100% health', to the 'left = 0% health'. All VAS-scales were anchored according to the following general rules: Patients were advised to rate their health status below 75% if they felt 'medical action is needed', they should rate the health status <50% in case of a 'strong need for medical action' and <25% in case of a 'medical emergency'.

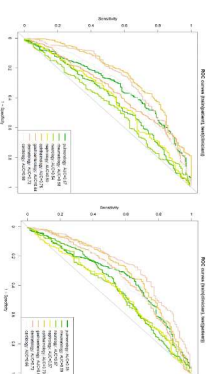
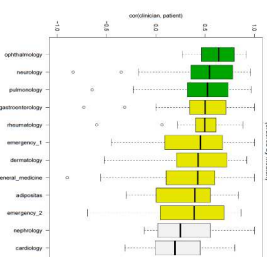
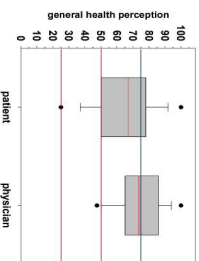
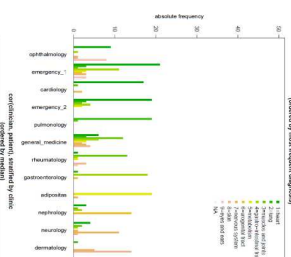
336 patients from different outpatient clinics (cardiologic, pneumologic, gastro-intestinal, nephrologic, neurologic, dermatologic, rheumatologic, ophthalmologic, and obesity outpatient clinic) as well as patients from internal emergency clinics and a general practitioner clinic were evaluated. Both, patients and the attending physicians completed the Pogen-OSSA. In addition the attending physician was asked to document ranking of the 5 most important diagnoses of the patient.



Example of OSSA results from two outpatient clinics

Statistical analysis was carried out using non-parametric testing. Furthermore, to predict main diagnoses based on patients' as well as physicians OSSA state-of-the-art machine learning tools, namely support vector machines (SVMs), were applied. To assess model performance multi-class AUC (area under the ROC curve) according to [Hand and Till (2001)] was estimated based on repeated cross validation (10 folds, 5 repeats), optimizing the SVMs hyperparameters using grid search.

Results: The test showed a good reproducibility. With a mean percentage of 66±17 SE and 74±0.98 SE, respectively, the physicians OSSA rating was significantly higher than the rating of the patients (p-value <0.001). Models predicting main diagnoses were constructed and estimated to perform with multi-class AUCs of 63.5% and 73.4%, based on patients and physician's OSSA, respectively.



Conclusion: In this preliminary trial with low sample size the Pogen-OSSA showed a good reproducibility and allowed a correct allocation of the patient's clinical problem to involved organ system by SVM analysis with multi-class AUC of up to 73.4%. These data merit further investigation of the Pogen-OSSA on larger patient cohorts. An APP-based version of the OSSA is developed.

References

David J. Hand and Robert J. Till (2001). A Simple Generation of the Area Under the ROC Curve for Multiple Class Classification Problems. *Machine Learning* 44(2), 171–186.

Corinna Cortes, Vladimir Vapnik (1995). Support-vector networks. *Machine Learning* 20(3):273–297

*©© Patent and Copyright pending